

Structure of everninomicin (Ziracin)

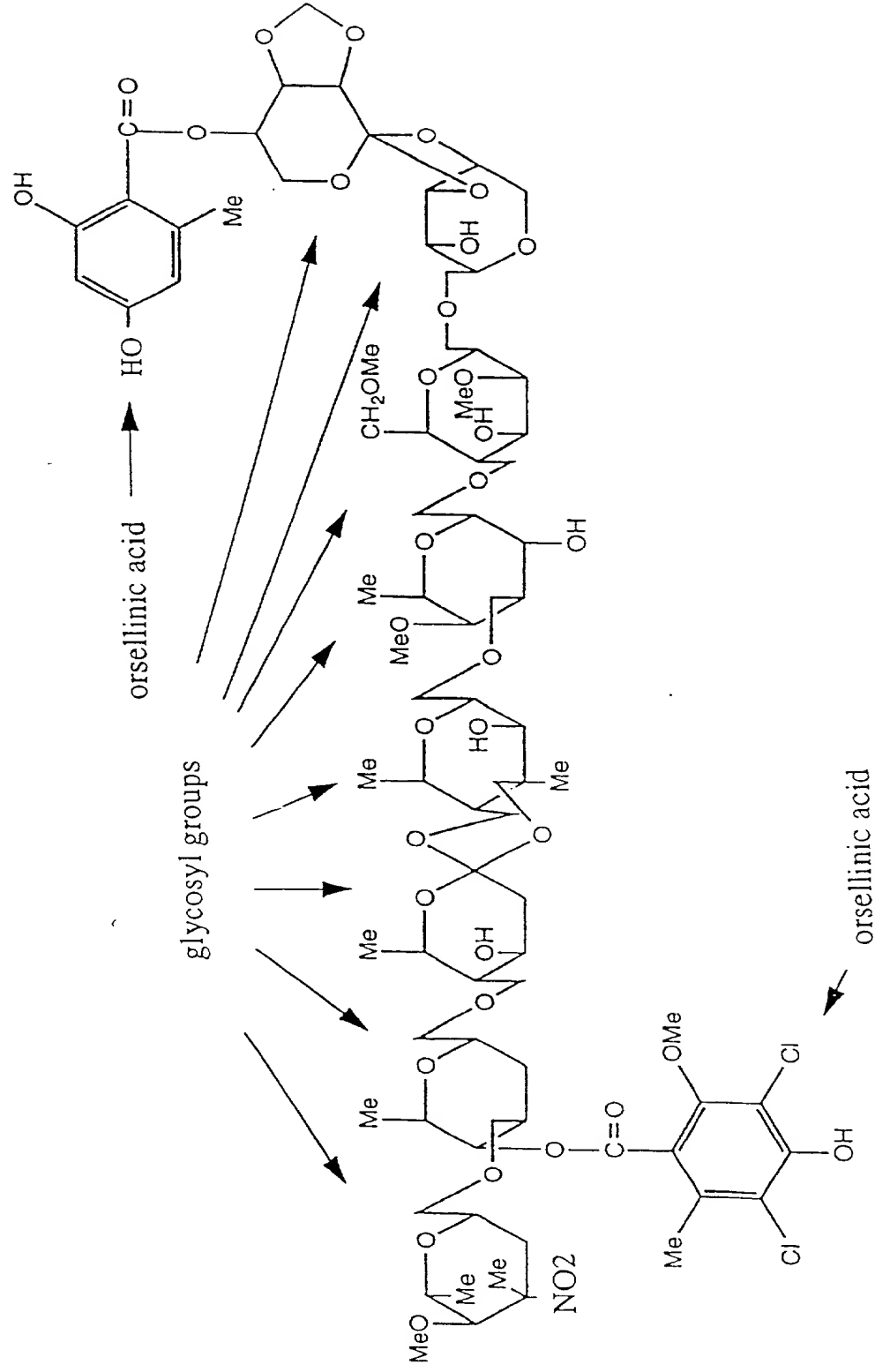


Figure 1

Overlapping cosmids and clones spanning 185 kb of chromosomal DNA containing the Everninomicin Pathway region.

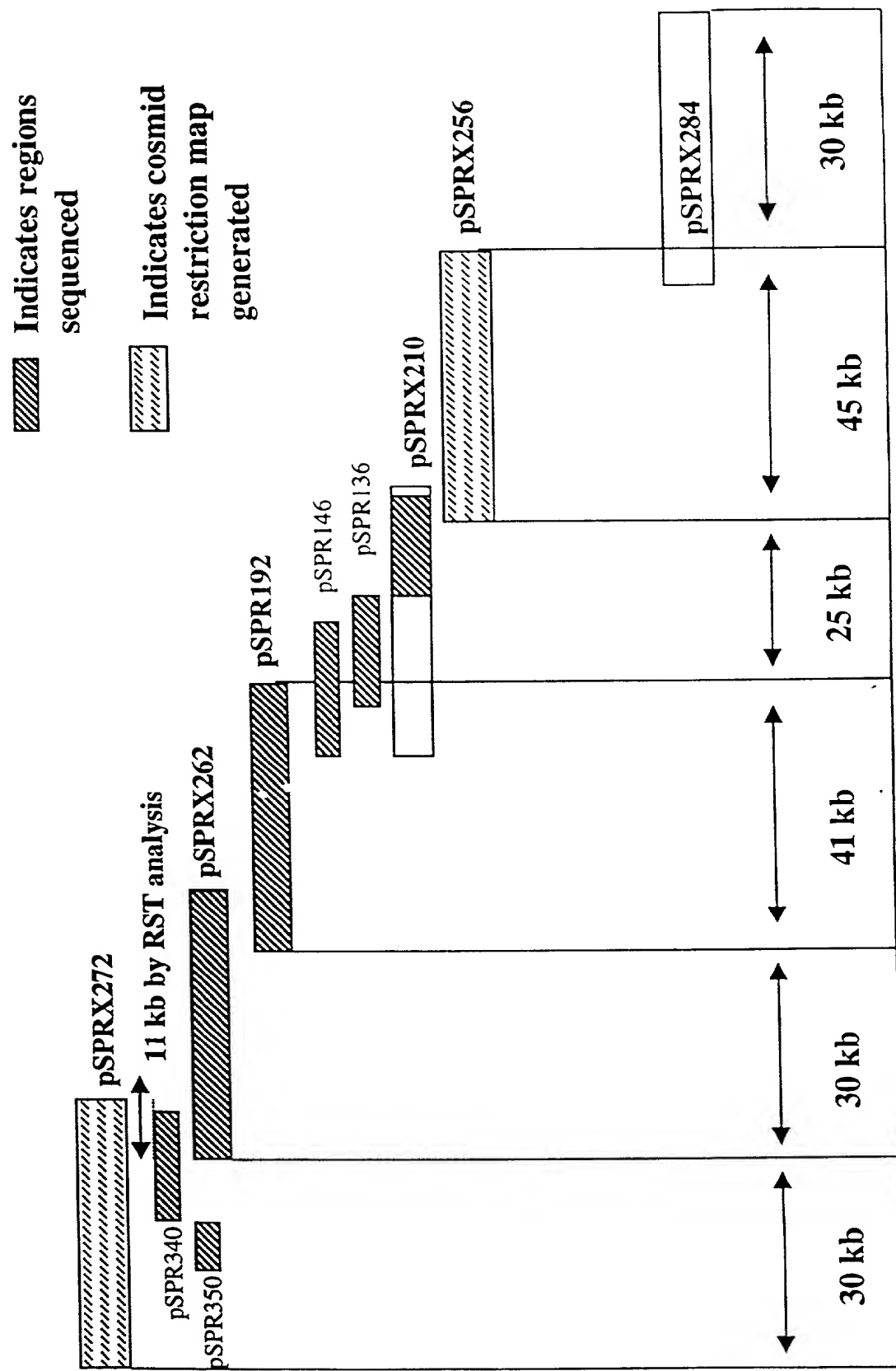


Figure 2A

pSPRX272
37.5kb

Cosmid pSPRX272

Regions sequenced indicated
by crosshatches.

Fragments cloned indicated
by clone designation beneath
fragment.

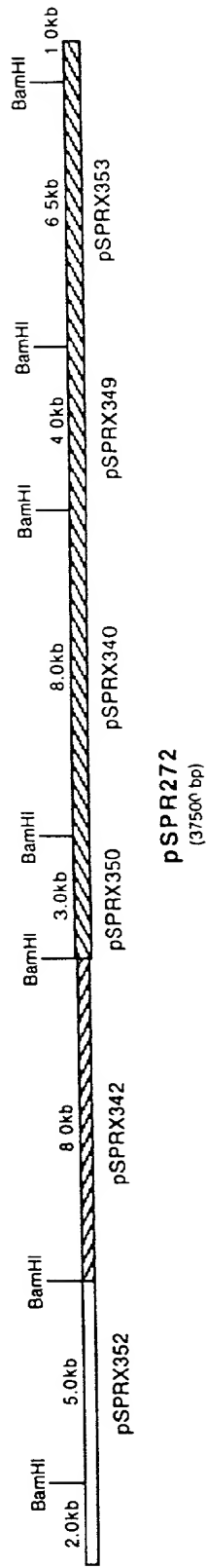


Figure 2B

Cosmid pSPRX256

Regions sequenced indicated
by crosshatches.

Fragments cloned indicated
by clone designation benign
fragment.

Cosmid pSPRX256

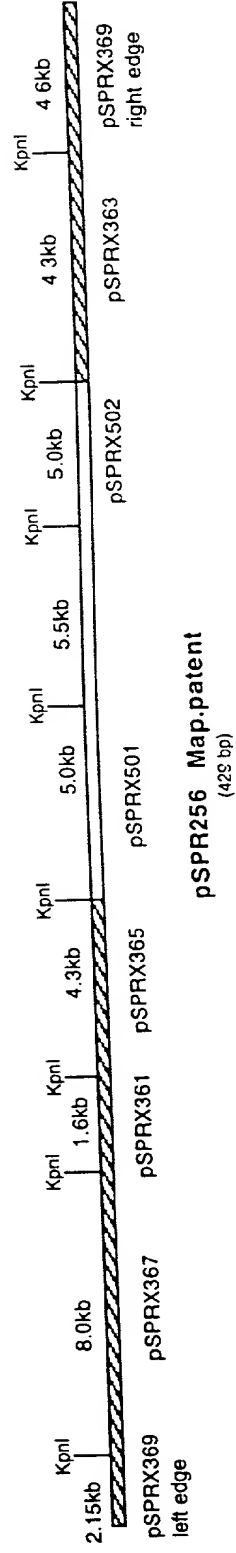


Figure 2C

Figure 3 (B)

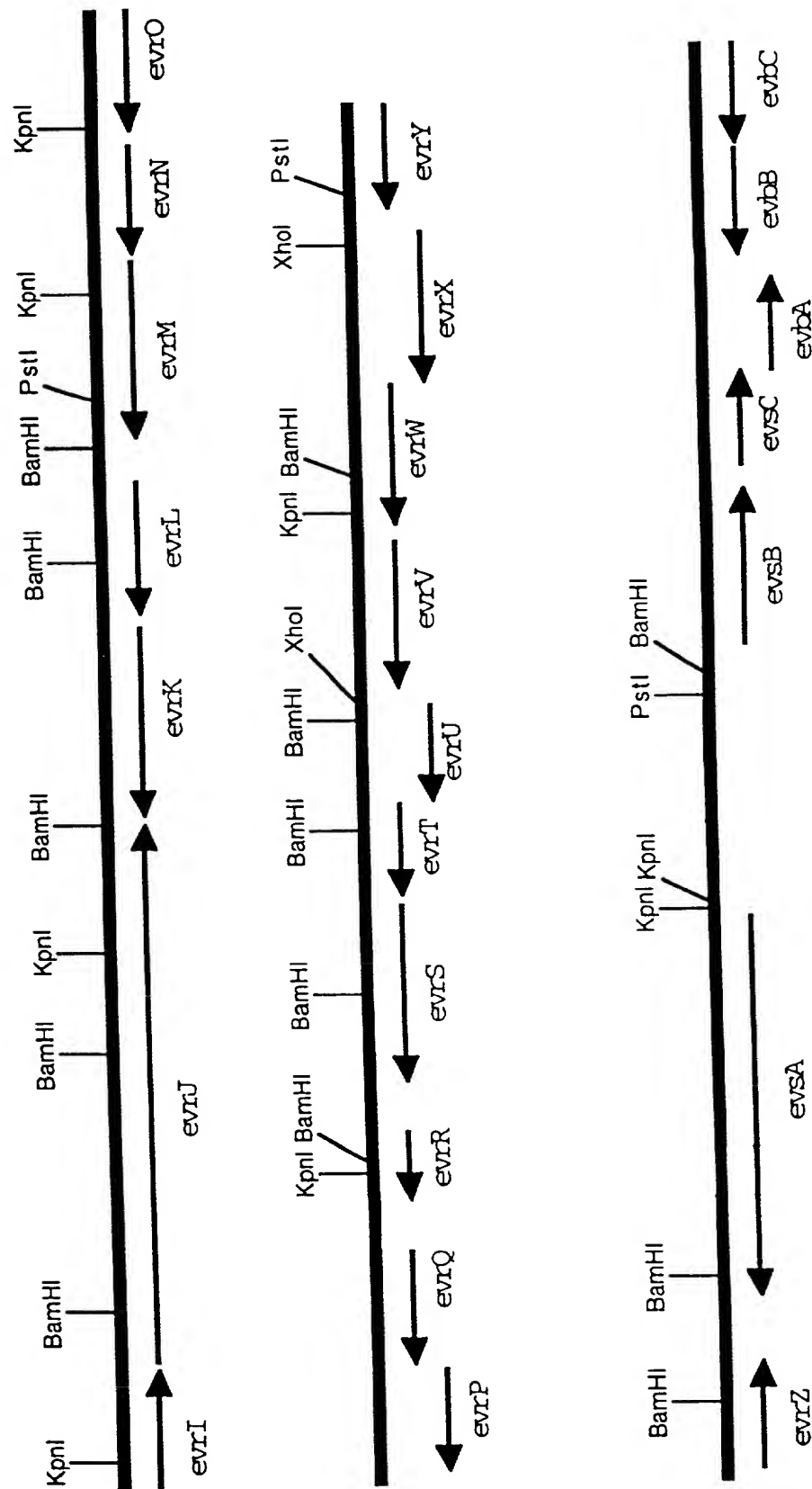


Figure 3 (C)

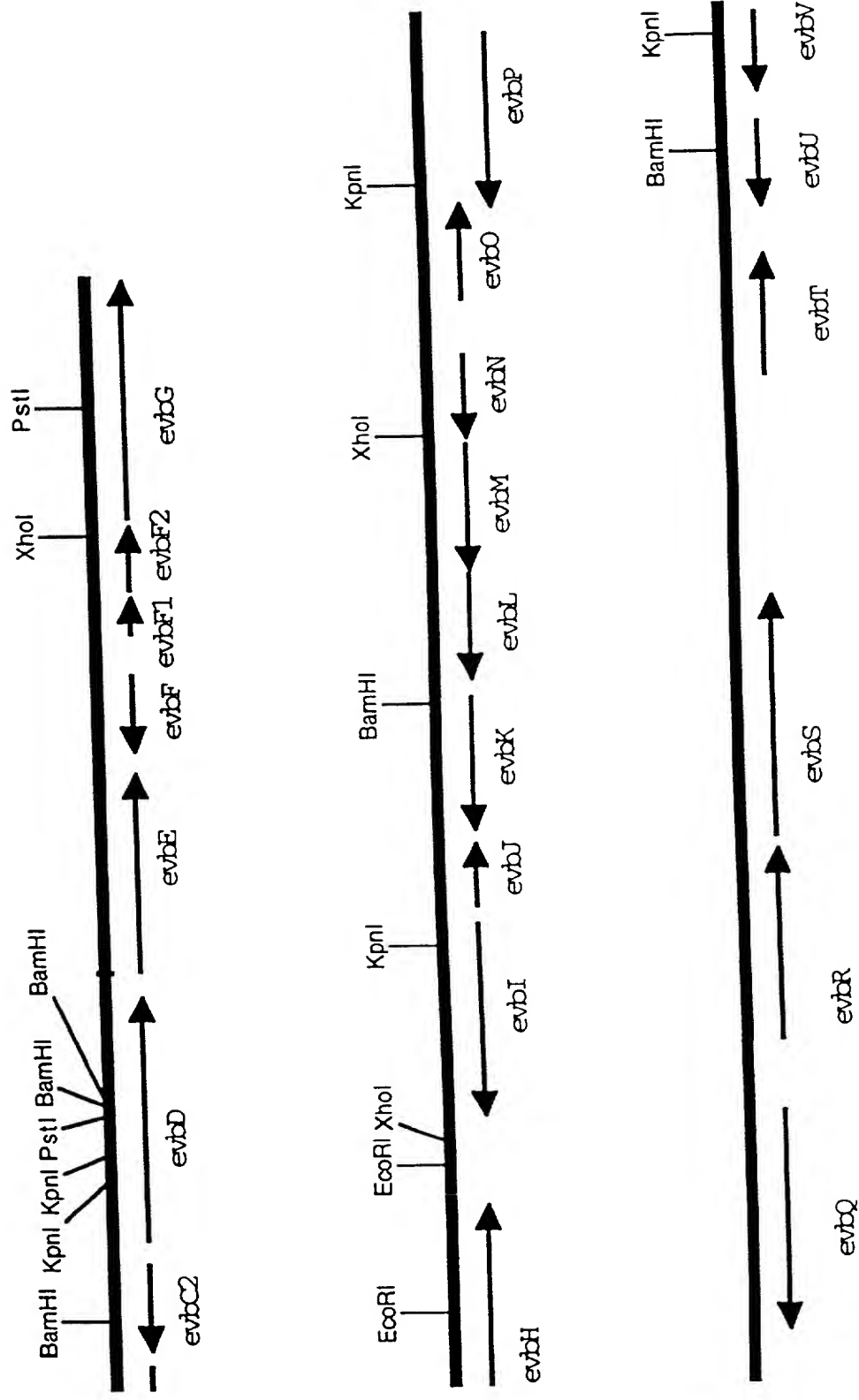


Figure 3 (D)

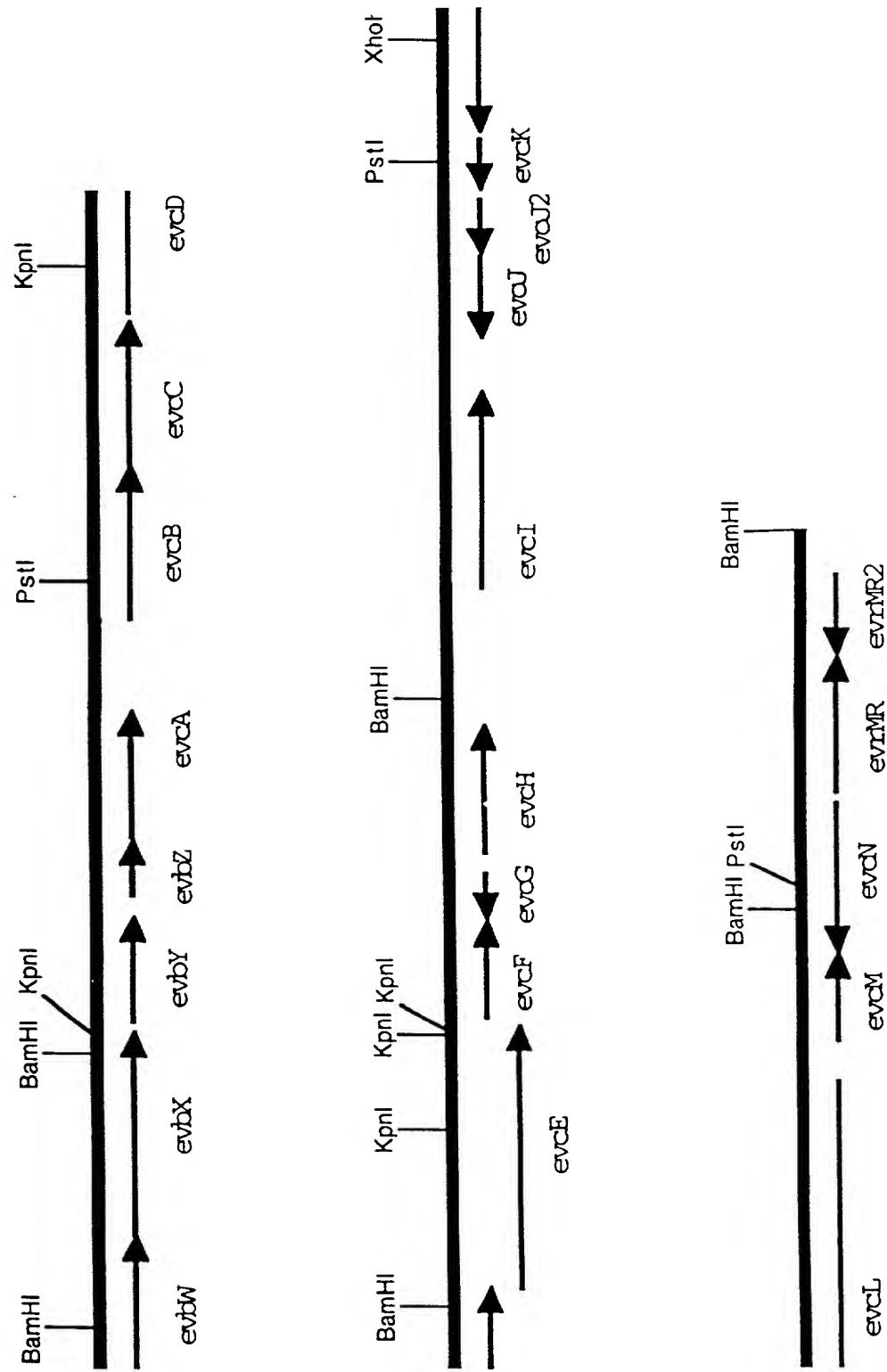


Figure 4 (B)

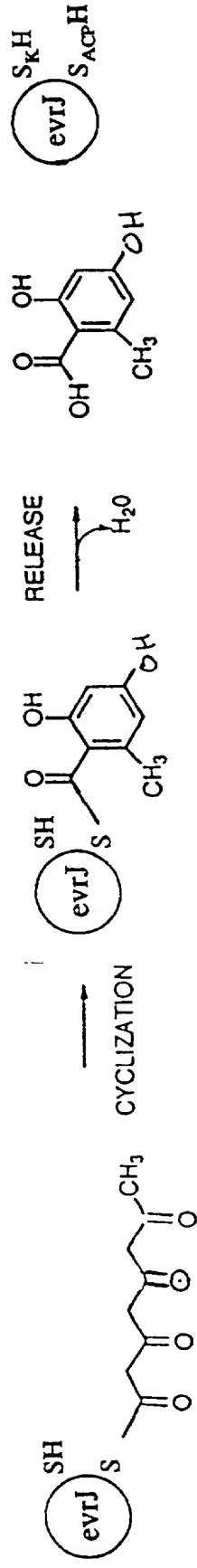


Figure 5 (A)

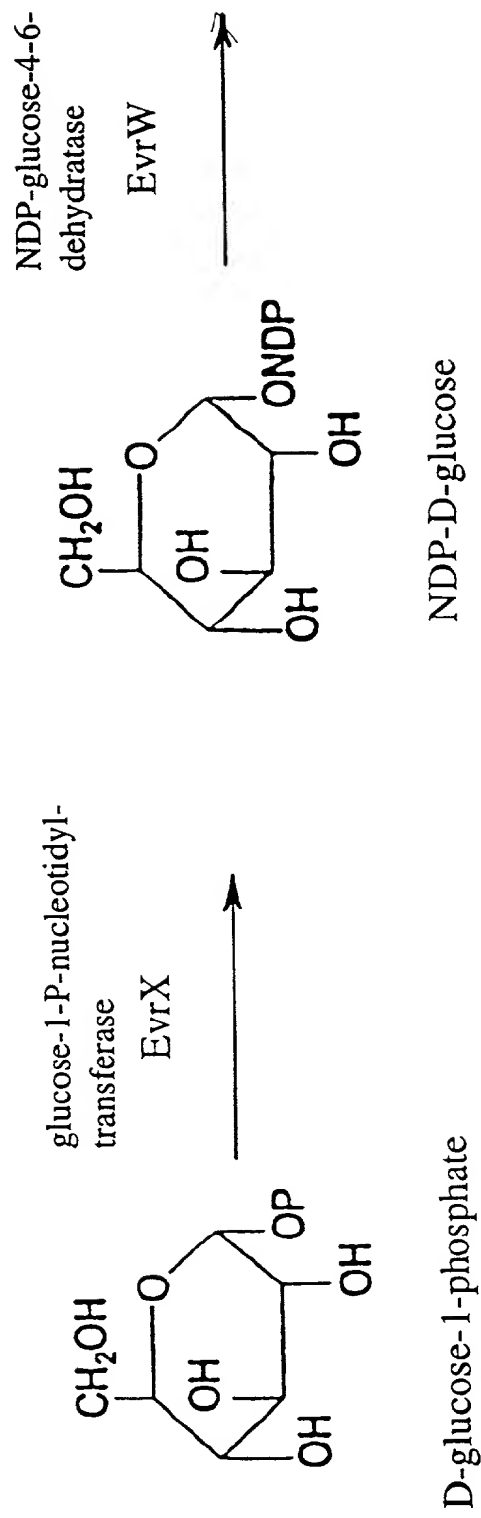
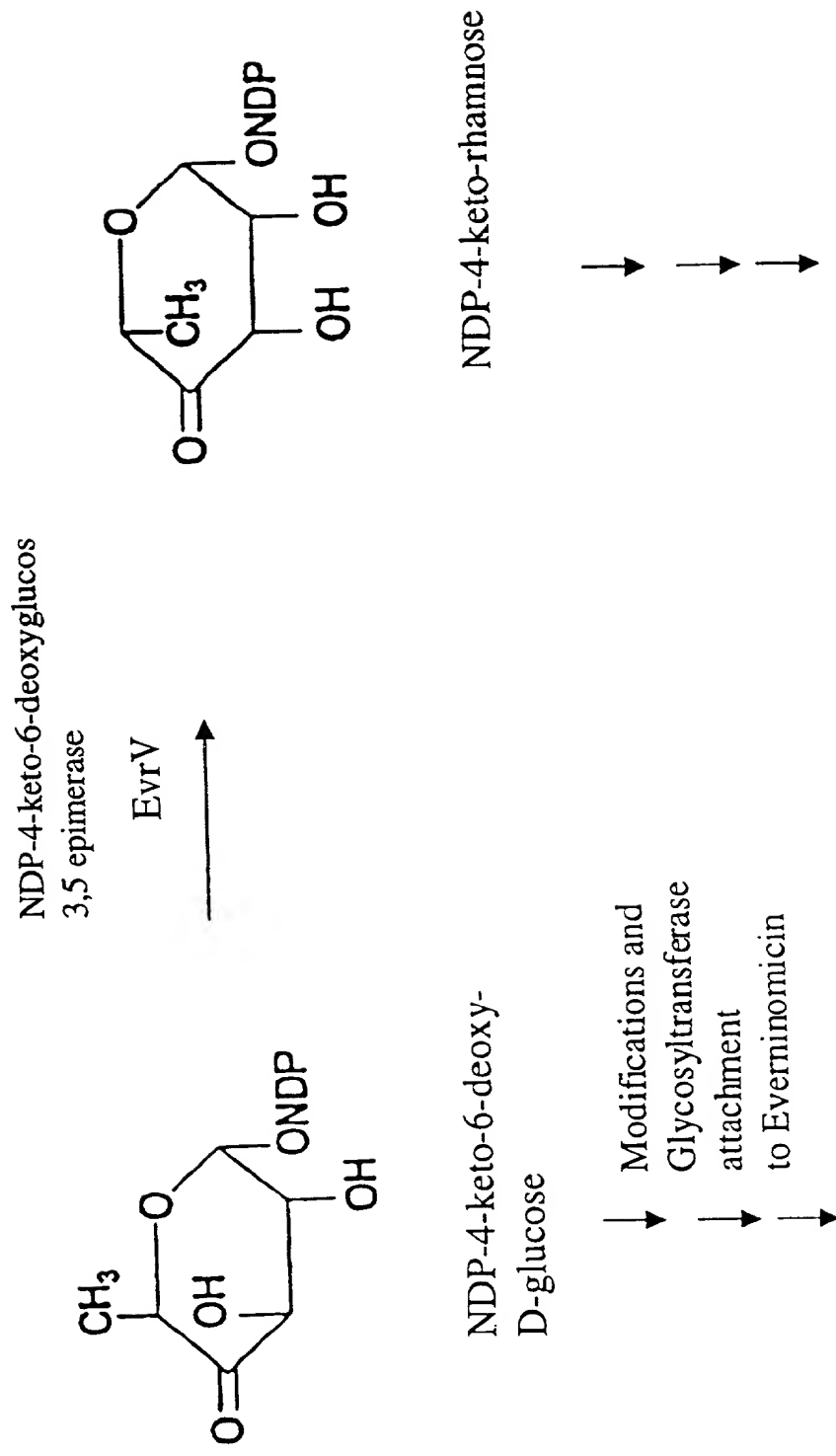
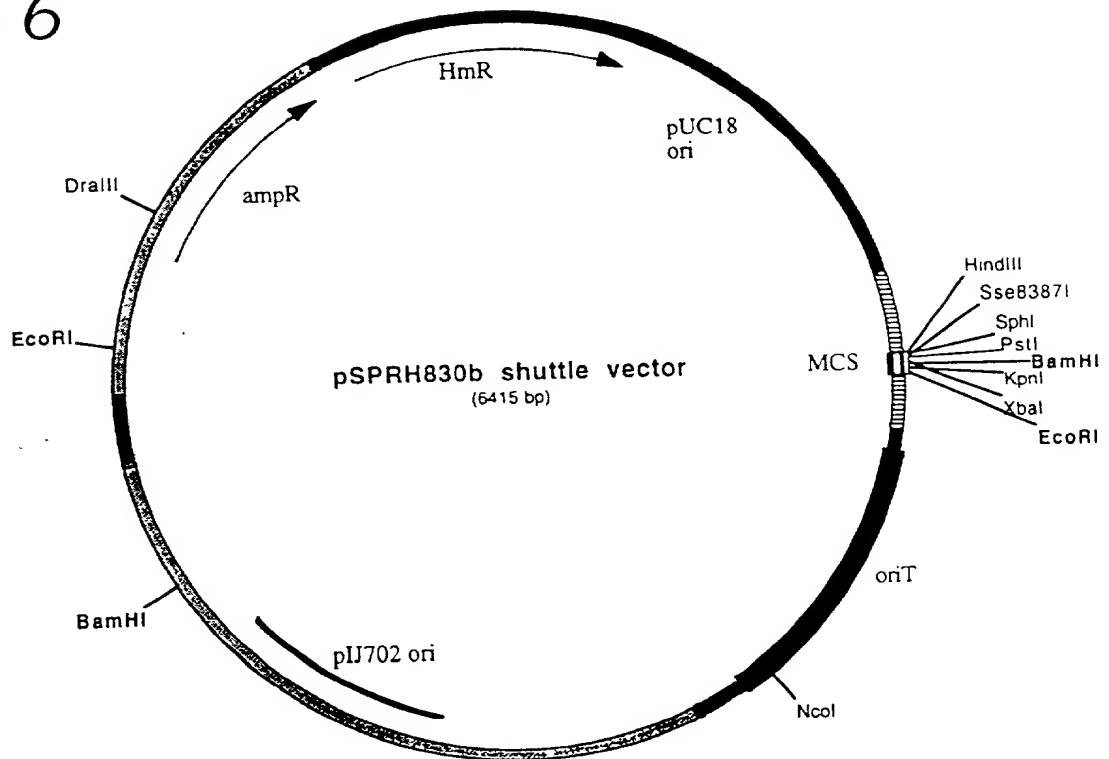


Figure 5 (B)



pSPRH830b *E.coli*-*Micromonospora* shuttle vector

Figure 6

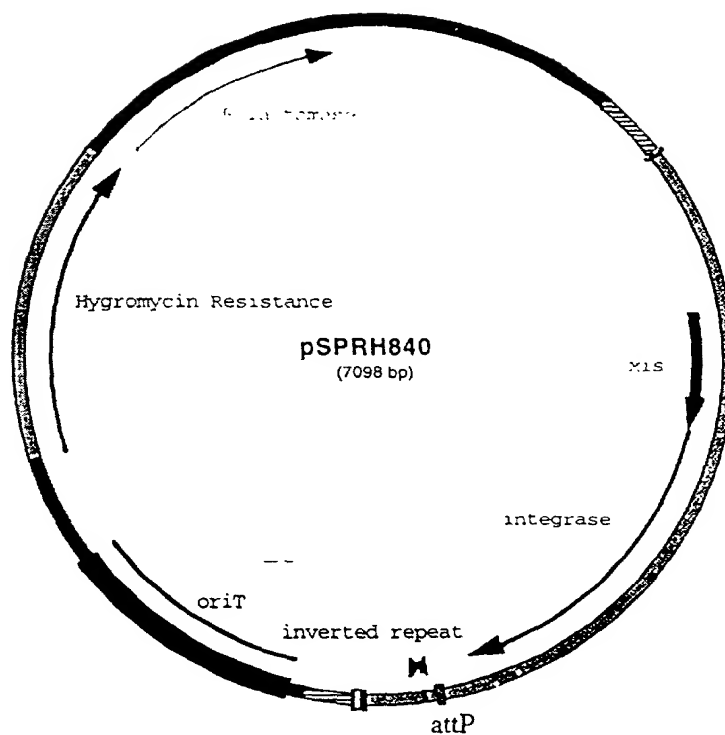


pSPRH830b - pSPRH826b backbone

Function	Source
- Ampicillin resistance	(pUC18)
- Multiple cloning site	(pUC18)
- pUC18 origin	(pUC18)
- Hygromycin resistance	(p16R1)
- oriT (origin of transfer)	(pRL1058)
- pIJ702 origin of replication	(pIJ702)

pSPRH840 integrating vector

Figure 7A



pSPRH840 - pSPRH826b backbone, pMLP1 *xis*, *int* attP insert

pSPRH840 conjugated
from *E. coli* into

HmR transformants obtained

M. carbonacea

+

M. rosaria

-

M. halophitica

+

pSPRH826b Insertion plasmid

Figure 8

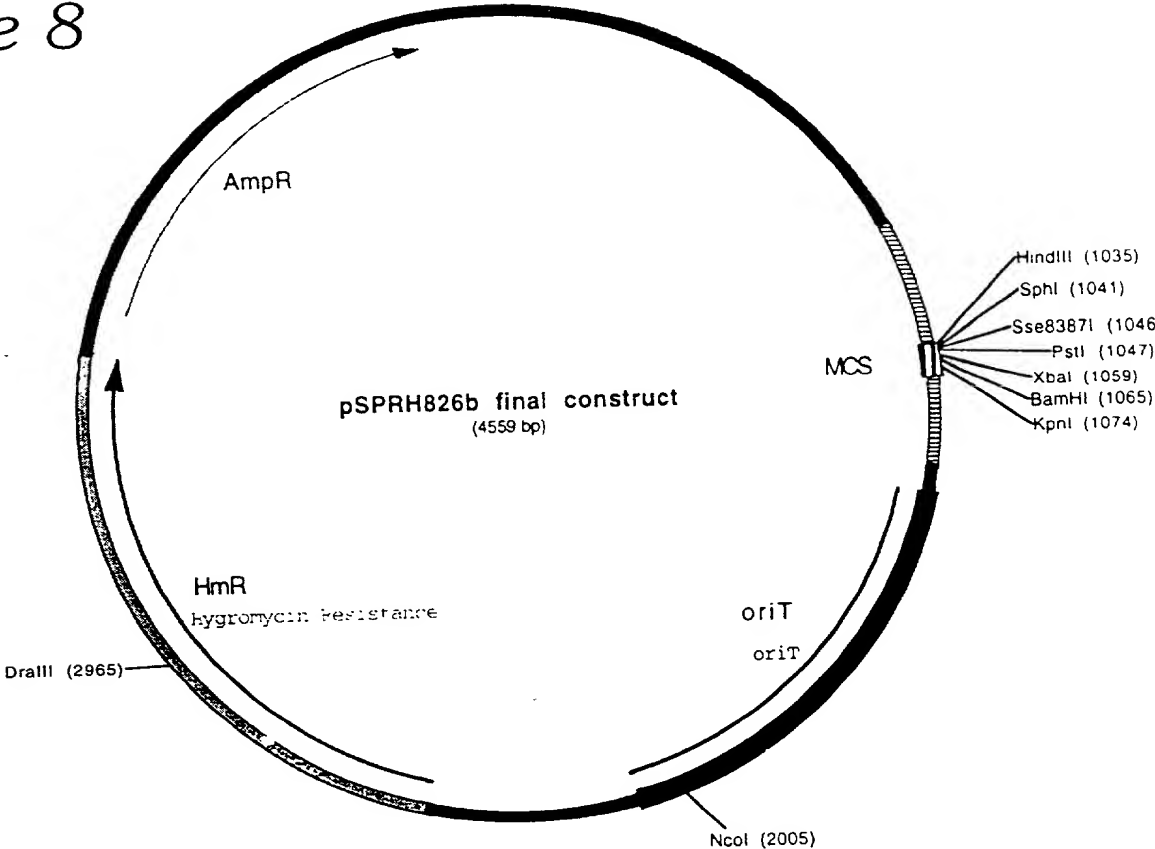


Figure 9

Analysis of *M. carbonacea* and *M. halophytica* pSPRH840 insertion site AttB/AttP region

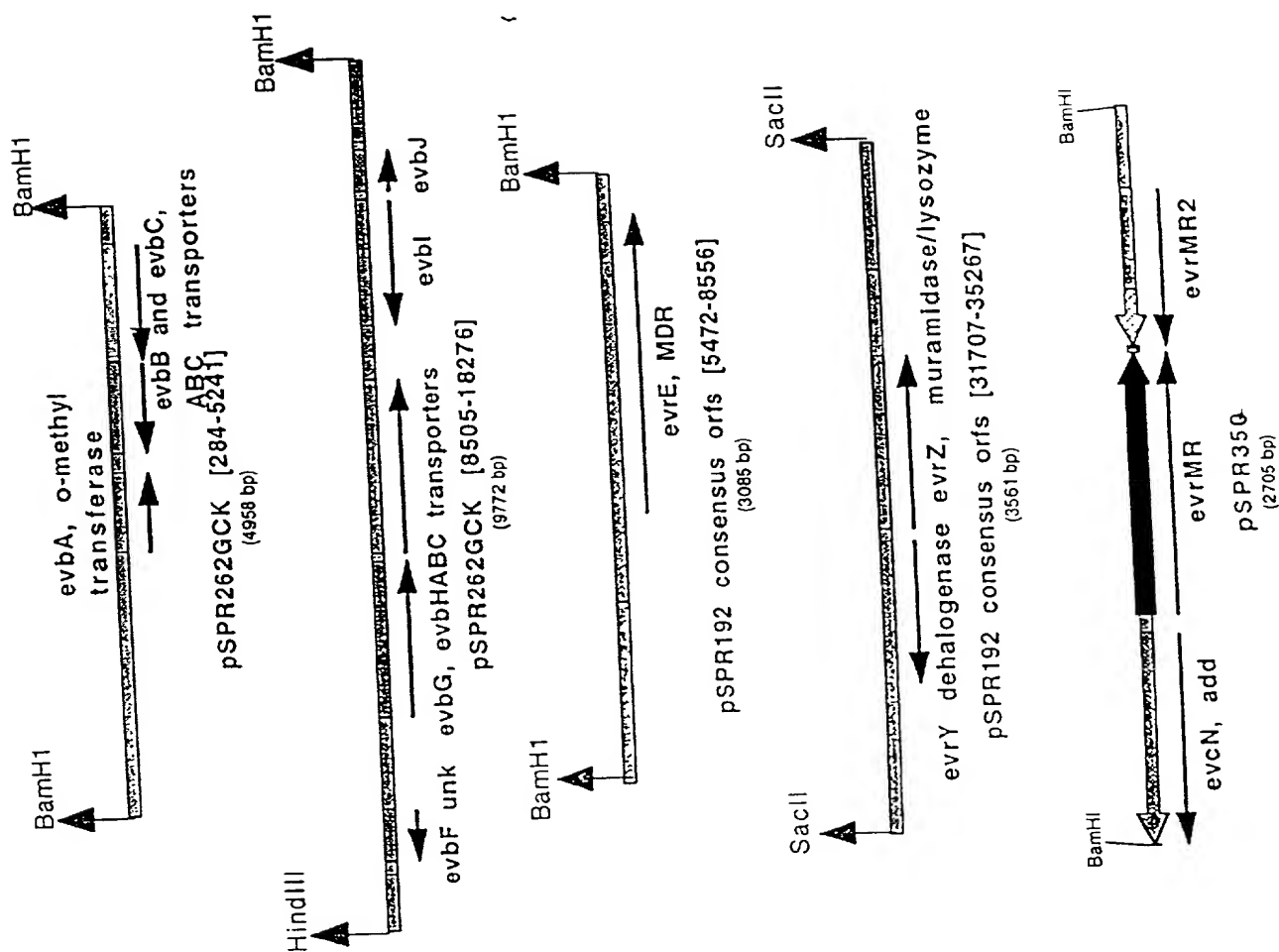
A. Alignment of pMLP1 attP region with religation clone edge sequence

<i>M. halophytica</i> PstI relig-9	60
<i>M. carb</i> PstI relig-1	59
<i>M. carb</i> PstI relig-4	60
pMLP1.intTCA.att region	60
Consensus	60
<i>M. halophytica</i> PstI relig-9	120
<i>M. carb</i> PstI relig-1	119
<i>M. carb</i> PstI relig-4	120
pMLP1.intTCA.att region	120
Consensus	120
<i>M. halophytica</i> PstI relig-9	179
<i>M. carb</i> PstI relig-1	174
<i>M. carb</i> PstI relig-4	175
pMLP1.intTCA.att region	175
Consensus	180
<i>M. halophytica</i> PstI relig-9	200
<i>M. carb</i> PstI relig-1	200
<i>M. carb</i> PstI relig-4	200
pMLP1.intTCA.att region	200
Consensus	209

Insertion junction

B. pMLP1 attP region

1	TCATCAACTCTAGGGAGGGGTAGGGGAATCCACTCCGGAGCCCTCCGAGCAATCCGGAGATGAGCGGAGCAACACAGCGTCAAGTGGCTT
94	GTTCACCCCTGACACAGGGCCCCCGGTA CGGGTTTCATATTCGCATCATGTCA CC CCGCAAGTGGATCTACTCCACAGCAGATC
174	AGCCCCCTCCGAGAGAGGGGGCTTCATGCTCATAGGGGACAGGTAGGGGAACTCA



< A A E A S I R P I V G A L L D V V T R G A P T L P A R L D A A
 3857 CCACCCATCTCTCCTCGGTACAGATTAGACATCGCCTGCTTCGGTTCGCGCTGTGCCGAACCTGTGCTATCAGGGTGC GCGCGATCACC
 < V
 3947 AATTGCTGGCTGATTGTCCACCGACGATGCTCGACAGGGATACCCAGAATAGGCGGCAACGGCTTGGCGAAACCCCTGTGCTGCGCGGAGT
 4039 TCGCCAAAGTCTAGTTGGATCACTCGAGCTTCCCCCGCAGAAGCCGTGAACCATGGGCCAGCGGTTGACGTGCGTATATATTCGGCCGACACG
 4131 TCGCGAGGACTCGTGAAGATACTGTTTCATCGCAGGACCGAGCAAGTCCAGCCTATTTCGGCCTGGCCCCACTGGCAATCGCGCCCGGATG
 > V K I L F I A G P T K S S L F G L A P L A I A A R M
 4221 AGCGGGCAGGAGTGTGATGGCTTCCACGCGAGGAGTGTACCGCGGACGATGCTCGGCTGGGCTGCGCGGCTTCCCGCTGGCGCGCTGAC
 > S G H E V V M A S T Q E V V P A T M S V G L P A F P L A A L T
 4313 CCTCGCCGAGCTCATGACCACCGACCGGGCCGCGATCCGCTGCGCATCCCGGCCGAGGACGCCCTTCTGCTCCCTTCTGCTGGCGCGGATGT
 > L A E L M T T D R A G D P L R I P A E D A A F V P F V G R M
 4405 TCGGCCGCTGGCGGATCAGCCTGGATCCGCTGCGCGACCTGGTGGCGGGTGGCGGCCGACCTGATGCTGCGCGGCCCGCACGCGCTAC
 > F G R L A A I S L D P L R D L V G G W R P D L I V G G P H A Y
 4497 GCGCGCGGATCTGGCCACCGAATTGGGGTGCCTGCGTGGCGCACCTGCTACCGGCAACCGGTTGGAGGCGAGGACCCATCCGGG
 > A A P I L A T E L G V P C V R H L L T G N P V D R E G T H P G
 4589 GGTGACGAGGAGCTGCGCGCGGAGCTGGCGCGCTGGCGCCAGGTGGCGCGCTTCCACCTGGCCCTGGACATCTTCCCGCCAGCA
 > V D E L R P E L A A L G L A Q V P P F H L A L D I F P A S
 4681 CCCGATCGACGAGTCCCGCGCGCGACCGGTGCGACCGCTGCGCTGGATTCCGACCAACCGAGCAGCGCGGTGGCGCGCTGGATGCTC
 > T R I D D V P P A Q P V R P L R W I P T N Q Q P V A P W M L
 4773 TCGCGCGCGCGCTGCTGCTGCTGCGCGCGCGCGCGGAGGTGGGTGGCGCCCTGCACGACGTGCGCGGTGGCGCGCGCGCGGT
 > S R G P R R R V L V T A G S L V T T T H N F D F L H G L A G T
 4865 CCTGGCGGAGCAGGACGTGAGGTGCTGCTGCGCGCGCGCGCGGAGGTGGGTGGCGCCCTGCACGACGTGCGCGGTGGCGCGCGCGGT
 > L A E Q D V E V V A L H D V P G V R H A G
 4957 GGCTCCGCTGGACGTGGTGTGCTGCGCCACTGTGACCTGATGCTGACCACTCGGACGATGACCGCGCTGACCGCTTGAACGCGGGGGT
 > W L P L D V V L P H C D L I V H S G T M T A L T A L N A G V
 5049 CCCAGATCGTGGCGGAGGAGCGGCTGCTGAGTGGCGCGCAACCTGTGACCTTGGCGGTGGCGGAGACCTCGCGCGCGGCGA
 > P Q L I V P Q E S R F I E W A R N L S T L G V A Q T L A P G E
 5141 GGACGCGCGGAGGCGGTGGGCAAGGTGCGCGCGCTGCTGCTGGAGGATCCGGTCCACGCGACCGCGCGATGCGCGCGCGGAGATCG
 > D T P E A V G K V A L E D P V H A T S A A A I A R E I
 5233 CCGAGATGCGCGCGCGCGGAGGTGCTGGCGCGCTGACCGAGTTCGCGACCGCGCGCGCTGACATGCGCGCTCTGTCACCGCGGAGCC
 > A E M P G P T E V V G Q L T E F A T R G L T C A S S
 > V T G G A
 5324 GGGTTCATCGGCTCCACCTACCGACGCGCTGCTCGAACGCGCGGACGAGCTACCGGTGCTCGACGACCTGTCCACCGGGCGCGCGAGCG
 > G F I G S H L T D A L L E R G D A G S V T V L D D L S R P E R
 5416 GCTGCGCGCGGGGTGCGGTCGACACCGGCTGATCGGACCGCGCGGGGTGACCGCGGTGGCGGAGCAGTGTGCGCGGAGGTCATCT
 > L P A G V P L H H G S I T D R A G L T R L A E Q C R P E V I
 5508 GCCACCTGGCGCGCGGAGCGGACGTGCGCAACTCGGTGGCGGACGCCACTCGGACACCGGGGTCAACGTGGTGGCGGACCGTCAACGTCTG
 > C H L A A Q A D V R N S V A D A T S D T G V N V V G T V N V L
 5600 GAGGCGCGCGCGGCGCATCGACCGCGGGTGGTCTTCCGCTCAGCGCGCGCGCGCTTCTACGGGAGGTGACGAGCTGCCCTCCCGGAGGA
 > E A A R A I D A R V V F A S S G G A L Y G E V D E L P S P E D
 5692 CGTCCGCGCGCGCGCGTGGCGCGCGGACGAGTACTGCGCGGAGCAGTACCTGGCGCTCTACAACCGGCTCTACGCTCGACCC
 > V R P A P W A P Y G A A K Y C A E Q Y L A L Y N R L Y G S T
 5784 ACGCGCGCTGCGGCTCGGCAACGTGTACGGGCGCGGACCGGAGCGGAGTTCGCTGCTCTACAACCGGCTCTACGCTCGACCC
 > H A A L R C L G N V Y G P R Q D P T G E A G V V S I F C G C L V
 5876 GCGCGCGCGCGCGGAGCGGTGTGCGGCGCGGCGGAGCGGAGCTACATCTACGTGGCGGACGTGGTGGAGCGGTTCCTGCTCGCGGT
 > A G R R V P D T W T P P G L V R G L V E F A Q R W P Q Y G V I
 5968 CGGCGCGGTGGCGCGCGGCTGTGGAACATCCGACCGGCGGACCTCCACGACGATCGGCAAACTACTGGACCTGGTGGCGCGCGCGCGCGG
 > G H G G P G L W N I G T G T S T S I R K L L D L V G R T A G
 6060 GCGTCCGCGACCGCGCTTCGAGCCACCGCGCTGGCGGAGTGAAGCACTCCGCGCTGGAGGTGACCGCGCGCGCGCGGAGCTGCGCTGG
 > R V P D P R F E P P R L G E L K H S A L E V T R A A R E L R W
 6152 GCGCGCGCAACGAGGCTCGCGGACGCGATCGCGAAGGTCTACAAGTGGGTGCGAGCGGACGAACCGGTCCGGGGGAGCGATGACCGCG
 > M T R
 > A A R T R L A D G I A K V Y K W V E A D E P V R G E R
 6242 AGGGGTCAACGCGCGCGGTAGGGTCCGACCATCGGTGCGGACCAACGAGATCGGCTGGCTGGACCGCGCTGCGCTGCTGCGC
 > E G S T P P V R V R V A L L D L G I Q H K G G G T A A S A
 6334 AGCGACAGGACCGGCTTCGAGCTGACGGTCTTCTACGTGGCAACCGCTCGGCGGACGCGAGCGTGGCGGACGTCATGTGCGGCTTCCCGG
 > S D T T G F E L T V F Y V D N A S A D G S V A H V M S A F P P G
 6426 CGTCCGGTTCATCGAAACCCCGCAATCTCGGCTTCACCGCGCGAACAACGTGGCATGCGGGCGGCGCTGGCGGAGGGTTCGACACCA
 > V R V I R N P R N L G F T G A N N V G M R A A L A E G F D H
 6518 TCTTCTGGTCAACCCGGACACCTGGACACCGCGGGGTGGTCCGCGGGTGGTGGAGTTCGCGGAGCGGTGGCGGAGGTCGCGCTTACCTGGCGGAGCAGCA
 > I F L V N P D T W T P P G L V R G L V E F A Q R W P Q Y G V I
 6610 GGCCCGTTCAGTACCGCTACGACCGCGGCTGACCGAGTGTGACCGACTTCAACGACTGGACGCGAGGTGCGCGCTTACCTGGCGGAGCAGCA
 > G P L Q Y R Y D P A S T E L T D F N D W T Q V A L Y L G E Q H
 6702 CACCTTCGCGCGGACCTGCTGGATCATCCCTCGCAGCTACCGCGGAGCGTCCGCGACCGCGCGCGCGCGCACCTGGAGCACGCGTACGTGC
 > T F A G D L L D H P S H V T A T V R D R A P R T L E H A Y V
 6794 AGGGCTCGGCGCTTCTGCTCGGGCGCGGCTGCTACGCGAGGTGCTGCGACGAGGTGTCCACACCTACTACGAGGAGGTGACCTG
 > Q G S A L F V R A A V L R E V G L L D E V F H T Y Y E E V D L
 6886 TGCGCGCGCGCGCGGTGGCGGGCTGGCGGGTGGCGGCTCTACTCGACCTCGGCATCCAGCACAAAGCGCGGTGGCACCGCGCGAGCGC
 > C R R A R W A G W R V A L L D L G I Q H K G G G G T A A S A
 6978 GTACAGCCGATACACATGCGCGCAACCGCTACTACTATCTGCTGACCGATGCGGAGTGGCGCGCGCGCGCAAGGCGCGCGCGCTCGCGCGC
 > Y S R I H M R R N R Y Y L L T D V D W P P A K A A R L A A
 7070 GCTGCTGTCTCCGAGTCCGTGGCGGGGCTGACGCGGTGGACGAGCGCGGGCGTGGGGCGCGGAGACCTTCTGGCGCTCGGGTGG
 > R W L F S D V R G R G V T G R T S A G V G A R E T F V A L G W
 7162 CTGGCGCGCGGAGCGCGGCTGATCCGGGAACGTGCTCGGCGGACCGGCTGCTGCGGCGACGAGGCGGCGGTGGCGCGCGCGCGGAGCGG
 > L A R Q V I R R A R R R L R A R G T G V D R A R E R
 7254 GAAGGAAACGTGCGGGGATGACGAGCCACGATTCTGCTCGCGGCAACTTCCACTGGCAGGCGGGTTCAGCCAGACCGTCCGCGCGT
 > K E T V R G
 > M S R P R I L V A G N F H W Q A G F S Q T V A A
 7345 ACGTGGCGCGCGCGCGGAGGCGGAGTGGCGGCTGCGCGCGCTGCTCGCGGCTGCGCGCGGTCGAGCGCGGAGCGCGCGGACCTGCGCGTC
 > Y V R A R E A D C E V R L C G P L S R V D A E T A R L P V
 7437 GAGCCGACCTCCGCTGGGGACCCACCTGGTATCATGTTGAGGCGCAAGCAGTTCCTACCGAGGCGCAACTGGACCTGCTGAGGCGGT
 > E P D L R W G T H L V I M F E A K Q F L T E A Q L D L V E A F

Figure 11B

Figure 11C

11475 T TCGCTCTGGCCAGCTCGGGCGGGGCGGTGTACACGCCGACGGTGTGGCGCCCTACCACGAGCGGTTCGCCACCGGGCGCCCTCGGCGTA
11567 CGCGCGGGCGAAGCTTCGGCTTGAACAGGAGCTGCTGCGCCACACCGGCGGTGACGCGGTGTTGACCGGCTGAGCAACGCTCTACGGTC
11659 CGGGCGAGCGCGCGACCCCGGTACGGTGTCTCTGTCACTGCTGGAGGCCACCGTGCGGCGAGAGCGCATCCGGCTTCTCGGCGATCCG
11751 GCGGCTGGTGGCGGACACGTACAGTGCAGACGTCACCGCGATCATGGAGGTCATCGCGCAGCGGGCGGTGACCGCGACCGGGACCGGCT
11843 CGCCACCGCTGTAAGCTCGGCTCGGGCTGCCACCTCCCTCGCCGAGTTGCTCCAGACGATGTCCACAGGTGGCGCGGTGCTGAGCTGGAAG
11935 TCATCCGGGAGCTTCGCCGGCAGTTGACCATTCGGGGCACTGGCTCGACACACCCCTCGCCCGGAGCACTTGGGCTGGCGCGCGGATC
12027 AGCCTCCCCGACGGCGTCCGCCAGTGTCTGGGAGGCGCTCTCACCCGGGCGGGCGGCCGGGGGTTCCCGGCCGACCGTACGCTCGGC
> S L P D G V R Q C W E A V L T R A G G P G G S P A R P S A R
< . G P
12118 TCGGAGAGCGCTCTCGGGGCGGGAACCGCGCAACCGCGCCCTTCGCGAGCAGTTCTGTTGGTCAACCCGCGCGGCGGTTCGCGCGGTGAGCC
> L G R A S R R G R E P P Q P R P S Q Q F V A Q P G G G A T R R G V A
< E P S R R P P F R R L R A R R L L E H V S L G A A T A A T V G
> .
12210 GAGGGCCAGTGGCAGGGCGAGCGGGATTCCGGGGCCGATCCGGTCCGGTCCCATTCCTTTCGCGAGCGCGCCCGGTGCTTGGTTCGGTCCGG
> E G Q W Q G E P G F R G R T V P V P F L A Q A G P C L V P S G
< L A L P L A L R S E P A T T R H R D W E K R L G A R A Q N R G A
> .
12302 CCTCGACCGCCCTGCCAGTACGCCCGCGCAGCAGGTACCGGGGTCAGCGCGCGCGGTGCGATGTATGGTGTGACC
> L A P P L P V R P P Q Q V P G G Q P A R V D I V M G D
< E C R G Q W Y A R R L L Y R P T L R G P D I D H T V
12382 GCGTGGTCCGGGAGCAGTTGCTTCGCGGGCGCGCGCGGCTTCATGGCGCTGATGAAGGAGGTGTCTCTC
> R V V R E Q L L A G A G G L H G A D E G G V L
< A H D P L L Q E R A G A A K M A S I F S T D E
12451 CTAAGTGGAGTTTCCCCCGGTACGGGTGAGGGCCAGGTGGAATCCAAACCGTGGGCGTTCGCGGAAC
> P
< G S Q L N G G T R S L A L D F D L G H A H A F
12520 GCGGAGTCCACCCCATGACGCGCGCCAGATCTTGATGTTTCCCTGGTACCGTGCAGCGCGACCGAGGTGGAAC TGGCCGA
< A S D V G M C A G W I K I N G Q D R H W G V L H F Q G S
> .
12603 GGTGACGTACCACGCGAGCGACGCGCGGGCGCGCAGCGCGGTTCGCGACCGTGCAGCGCGCTTCGCGCAGGCTTTCGCGGACGCGTTCGA
< T V Y V P L R L A P R A L R T G V V H A G D R L S K R V A D V
12695 CGCGACCGGCGTCGAGCGCGCAGCTCGTCTGTCAGCAACATCAGATGGTGTGCGCGCAGCGCGGCGAGCATCCGGTTCGGGAGGCGCGACAGG
< A A A D G D L R V D D D V F M L H H H P W A R L C L M A N R S A
12787 CCATTTGGTGACCGAGGATGCGCATGTTGCGCGCGCGCGGACCTCTCGCGGACCTCTCCGCGTGCAGCGGCGGTGACGGCGGTTCAG
< G N T A G L I R M T G G A A R V E A A V E A E A T V P R D L
12879 CAGGACGTAGTACTCTGCGCGGAGAGCTGGGCGCATCTGTGCGCGCAGGTGTTTCTGACGTTCTCCACCGGAAACCGCGAGATCGCCACCA
< L V Y Y E D G S L Q A M N H A L H K R V N E V R F A C I A V V
12971 CCATCGGGTGGTTCGACCGGATCGCGGCTGACCCAGAGGCGTGTGTTCCGGCATCTGCTCTCGACATGAGCGTTGTGGGCGCGCGGGGAGA
< M P H D S P D R S V S A N N P M
13061 GCGCGCCACCGGATCGCGCTCCGCGGCGAGCGTGC CGCGCGCCCGCGCGCGCCACCGGACAGGTGGCGCACCGCGAGCGCGGCGGGCGTCCGG
13153 TGAGCGCGCGGGGGCGGAGCTCAGGCGCGTCTGCTCTTGTGACGCTCGCGCGCGCAGGTGCCCGCGCGGAGGATGCCCATCGCGTACGCGCT
13245 CGCGATCAGGCGCGCGCGGTGGCGCACTCGCAGGTTCTGCAACAGCTTCTGTCAGCTGAGTCTCGACGCCCTGACGGCTGAGGTAGACCTTGT
13337 TCGGCGATATGCAACCGTCCGTTTCGCTCGCGGATCTTTCGATGTGCGCGCTCGCATCGCAGCTGAGCACTGAGAACTTGAGATCAACAGCGCC
13429 CCTTGATTTGTGATCTGTTGGCTCGTACCGCGCGCAGCGTACGCTCGCGCGGACACCCCTGAGGTGCGCGGTGAGGAGGTGTTGCTTCCGCT
13521 GGAAGACGCTCAGCGCGGAGCGCGGTTGCCCGCAGCTCTGCTCTGCGCTTGGGAGCTGTCTGATGTCGCGGCGATGCCCGCGCGACG
13613 TAGCCCCGCTGACGGGCGTGTGCGCTTGGGCTCCAACCTTCGGCAACCCATCGGTTTCGCGCTGCAAGAACAGCGGAGGAACATTGAGTTTC
13705 GAAGCCGCTTCGCTACTGTCGATCGACATCACTGGCGAAATGAACCCGACCGGATTCGAGCGCGATAGTTACGGAAGTGAACATTCGGCTTGGC
13797 CTGTCGCTCGCGCTGACTACCAATACCGCTAAGCTGAGCGCGCGCGCTGCGGAGGCTGACCCCGCGGAGGATTCGCCAGGCTTCCAT
13889 CAGAAGTTCGCGAGGATCTTTCATCTCAGCGCGACCTCGGCGCAAAACCCCTGCTCAAGACCATGAGTAAAGCAGCGCGGGAATCCATGCGAT
13981 GACATGTGTCTTTCAGCAACCCAGCTCCAGCCGACCCACCCCTGACAAAAGGGCGGAATTCGCGACCGAGCGCACCCAGCACATTC
14073 CTAGGGGATCTCTAGTCTGGCGAGCGTGCATCCCCACCCGCTCAATCAATTGCGCGAACAAATGCGCCGCAAGATCAACCATCTCG
14165 CACCGTCAGGAGCGGCGCGCTCGCGCAATTCGGAACCGAGCGCTCACCCTGCGGTGAGCTTCAACCGCGTTCGGGTGAGCTTCAACCGCGCTGGGGGGGACCGGCGCGGA
14257 CACCGTCGCGCTGATCAGCGGTCTGCTCGCGGCTGTCGCGCGCGGCGGACCGCATCGCGCGGGGCGCGGCGGAGCAAGATCTCG
14349 GCGACCTGTCAGCGCTCGAAGCGGACCGCATCGCGCGCGGGGCGGGAACACCCGACGCTTTCAGCGCGCGGACCGATCGGGAAGCGGACT
< . R R S R D P L R S
14440 CGCGGCTCTGTCAGAGCGCGCCACCGCATTCGCGAGCGATCTGTCGGGCGCGCCACCCGAGCAGGTTCACGCGCGGGCGCCGGTTCGACGC
< A A E D W L G G V A D R L S R R P R P R W G L L D R A P G P D V R
14532 GGGCCAGTTCGACGACCTCCACCGGACCGGGCGGCTCCGCGCACTCACCACCTCGGTTCGCGACCCCGCTCAGCTCGATCAACATCTCCACG
< A W D V V E V S G P R D P L E V V E T P V G S V E I L M E V
14624 AGGAGCGGACCGGAACCGGCTCCCCCGCGCCACCGCGATACCCGCGCGGTGACGCGCTCCGCGTGGCGGACACCCGCGCTCGGCGAC
< L S R V P V A E G R G V G I V R G T V S E R T A S V V A E
14716 GTCCCGCAGCTCGACGTAGTCCCGTGGGCGCGCAACCGGAGCAACTCACCTTGGCGGACCGGTTCGCGGCGCGCGCTCGACAGGCGAA
< D R V D V Y D R H A R L P S L E V R A S R D R G A A D V L R V
14808 CGAGCGCGCGCCCAACAGGCTGTGCGGTGGCAGCGCCGAGCCACCGGATTCGCGAGCGCGCAGCAGCGTTCGCGTCCAGCGAGCGCGCGG
< V V R G L L S D P P V G P G V V N A L R L V T A D V S G A R
14900 GTGCGCGCGAGCAGCGGCTGGGTGGGCGGAGCTTCGCGCTGCGGTACATGCTCTCCGCTGGGTGGGCGGACCGTGGCAGCGGCGGAGCCCC
< T A A L V A Q T A A L K A R A G Y M S E P Q T P V T A G A P A G
14992 GGGTGGCTCTGAGCGGTTCCAGGACCGAGCGGAGTGCACAGCTTGGCGCGGCAACCGGTTCGCTTCAGCGCGCGGCTGACACTCCAGC
< P P E Q V R E L V S G L H V L R P R C R T R E L A A T V S W T
15084 TCGGGATCTGTCAGCTCGACGCGAGCTCGCGGCTCGCTAGGTTTCACTTTCGCGCGGTGGCGTTGACCAACCGCTCGGGCTGCTCGGCGTGC
< P I T C S S P L E A D T L N W K G G T A N V V A D P Q E A
15176 AACACGCGCGCGCGCGGCGCTCCAGGTGGCGAGCTTCACGCGCGCGGCGGTCAGCGACGCCCGCGCGGACCGGAGCGCGGCGGCGGCA
< F V A A L A A P E L T A V A D L A R A R Y P L G A S P V R R A L
15268 CACGAGGACGTCTGTCGCGCGCGCGGCGGCAAGCGCGCTCAGCTGACGCGCGGACGAGCGCGTACCACCGGACGACGACGACGACGCGCGCG
< V L D V D G R A A L A A S V H R G V F G T G G V V V R R A G
15360 CCATCCGTACTCTTGGGATCAGTCTGTCGCGCGCGCGCTTCAGCGGACCGCGCGCGCTTACAGGTACGCGGCGCGGCGCAAC
< M
< . D R A G A A D L R G G P G Q C T V P A R L V
15449 CTGGCGCGGTTCGCGAACCCTCTGTCGACGAGCGCGGCGCGGCGGATTCGCGCTCACTCTCTCCGCGCGGTGACCGCGGCTGACGCGCGGTA
< R A R D R V E D L L R A R A R I A T V E E A P Q G A C T A R V
15541 CGAATCTCGCATCTGTGTGAGCAACTGGTCTCGCGCGGGAAGTCAGCTCCGCGCTCTGCTCTGCGCTCCACCGCAGCAGCGGCTGC
< F E R M T N V F Q D E A P F T L E R T E D Q R E V R V P H
15633 CAGGCGCGCGGTTGGGTGTACGCGCGGTCGACGACGATCGCGCGCGGCTTCCCGAGAGCTTGGTACTCGACCGGTAGGAGTGTCTGGAAC
< W A P P P T Y A R D V I R G A S G W L Q Y E C R Y S H E F
> .

Figure 11D

> A T L V V F L A G S L L S G M A Q S I T Q L T V F R A V H G L
19676 CGGCGCGGGCGGCCTGATGGTCTGCGCGCTTCGCCATCATGGTGAAGGTCTCGCCGCCCTCAGACTGCCCAAGTACCAAGGGGATCATGTCCG
> G A G G L M V C A F A I M V E V L A G P D L P K Y Q G I M S
19768 CGACACTGGGCTGACCATTGGCGGCGCCGCTCGTCGGCGGCTGATCACCGATGAGCTCGGCTGGCGCTGGTGCTTCTACATCAACCTG
> A T M G L T M V A G P L V G G L I T D E L G W R W C F Y I N L
19860 CCGATCGGGCGGTTCGCTCTCATGTGGTGTGATGACCTGCGCGCCAGACACAAGGCGCGGATCGAATACCGGGGTGCTGCG
> P I G A V A L L I V V L M M H L P R R H T K A R I D Y A G A A
19952 CCTGCTCACCGTGGTCAGTTCTGCTGCTGCTGGTGAACCACTTGGGGCGGCATCACCTACCTGGGCGCTCTCCGATGATCTCGGGCTGG
> L T T V V S S C V V L V T T W G G I T Y P W A S P M I L G L
20044 TCGCGCTCGGGGTGCTGACCTGCGCGCTCTTCGTGGTGGTGCAGCGACGGGTGGCCGAGCGGTTGGTGCCCTGGCCATCTCTCCGAGCGCTG
> V A L G G L T C A L F V V V E R R V A E P L V P L A M F R S L
20136 AACCTTCACTGACACCCCTCATGCTCTTCGTGTCGGCTTCGCGCTCATCGCGGGCTGACCTTCCTGGCCCTGTTCCAGAGGCGGTGCA
> N F T L S T L I A F L V G F A L I A G L T F L A L F Q Q A V Q
20228 GGGTGCCTCGCGCTCCGACTCCGGCTGTGTGCTGCTGCCCTGCTGCTCATGGCGCGGTCAACGTGTGCGGGGTGCGCTGATGAGCG
> G A S A S D S G L L L P L L L S M A A V N V V G R L M S
20320 CGGGCGTTCCTACCGGCTGCTGATGCTCGCGCGTGGCGGCTGATGACCTTAGGCTGCTGCTCTTCGCGCTGATGACGATGGGCGACCG
> G G R S Y R L L M L A G A A L M T L S L L L F A L M D V G T S
20412 CGGACGGTCAACCGGATCCCCATGGTTCGGCTTCGGCGCAGGGCTGGGCTGCTCATGCAGACAGCGCTGATGTTGGCGCTGAGCAGCGTGA
> R T V T A I P M V G F G A G L G L M Q T S L M V A L S S V E
20504 GATGAGGAACCTCGGGGCTGGCGCTGACGCTCCAGCTCCAGCTTCCTCGCACCATCGGTGGGCGGTGGGGCGTGGCGGACGGTCTCGCTGTTCT
> M R N L G V A A S T S T L F R T I G G A V G A S A T V S S L F
20596 CCGTGGCGGTGACGTGCGCGCTGGCCGATCGGGGGTTCGCCGACGTGGCTGACCTCCTCGGCCACTCCGCGCGCTGGACGCGCGCGGCTG
> S V R V Q S A L A D R G V A D V A D L L G H S A R L D A A G L
20688 GCCCAACTCCCCCGGGCGCTCCGTGTCACCTTCATGACGCGGTGGCGCTCCGGCACCAGCTGGGCTTCCTGATGACCGTGTGGCGGGGCT
> A Q L P R A V R V H F M H A V A S G T R W A P L M T V L A G L
20780 GATCTGCGTTCGGCGCGGCTGGTTCCTGCGCGGGTCAACCGTTGACGTGGCACCAGTGGCACCAGAACCGCGCGCGACGCTGCGCGCGC
> I C V A A A W F L R R V T T L T G T A P V A P A P E P A R D V A A
20872 CCGCGCGCAGACGCGGGCGCGCGCAACTACTAGCGATTTCCTAGGTTCTCTGCTGCAGCGTAGAGCTGAATTCACCGCGACCTAAACA
> P A A S S G R A P N Y .
20963 TTCTTTTCGGATCCGAATCCGTCATTCCTCTGCTGGCGATGGTGCAGCGGGCGGGCGCTGGCGAGCGGACAGACAGATTCTCGGAT
21055 TTGAGCTCGATGTCGCAAGATCTAGTCACTGCTGGAGGTTCGGCGCGGATCCACGGCGCGCGCTGCTCGCCGATCGGGGCTGTGCG
> M S S K I L V I G I G G P A G S T A A L L A R S S G L S
21145 GTGACGCTTGGAAAAGGAGACGTTCCCGGATACCACTCGGCGAGTTCGCTGCTCTCGCGCACCATCGTCGATTTCTGGGCGC
> V T L L E K E T F P R Y H I G E S I A S S C R T I V D F V G A
21237 TCTCGACGAGGTTCGACTCGCGGGGCTACCCGCAAGAAGCGGGTCTCTGCTGCGCTGGGGCAACGAGGACTGGGCGCATCGACTGGGCAAGA
> L D E V D S R G Y P Q K N G V L L R W G N E D W A I D W A K
21329 TCTTCGGTTCGGGCGTGGCTCTGGCAGGTTCGACCGGACGACTTCGACCACTGCTGCTCAACAACCGCGGCAAGCAGGGCGCCAAGATC
> I F G P G V R S G V D R D D F D H V L L N N A G K Q G A K I
21421 ATCCAGGGCGCGGTGCAAGCGGCTGTTGTTTCGACGCTGAGCGGGCCACCGCCGAGTGGTTCGACCCAGTTCGGTGGGTGAGGTCGCGAC
> I O G A A V K R V L F D G E R A T A A E W F D P E S G E V R T
21513 CATCGATTTCGACTACGTGGTTCGACGCTTCGGCGCGGGCGGGCTGATCCCGTCCAGCACTCAAGCACCGGCGCCCGCCAGCAGCTTCA
> I D F D Y V V D A S G R A G L I P S Q H F K H R R P T E T F
21605 AGAAGCTGGCCATCTGGGGCTACTGGCAGGGTGGCTGCTGCTGCGCACTCTCCCTCGCGGGGATCAAGCTCATCTCCGCGCCGACCGGC
> K N V A I W G Y W Q G S L L P N S P S G G I N V I S A P D G
21697 TGGTACTGGGTCAATCCGCTGCGCGGCGACCGGTACAGCATCGGCTTCGCTGCGCACAGAGCGCTTCTCTGGAGCGGCGCAAGGAGCAGC
> W Y W V I P L R G D R Y S I G F V C H Q S R F L E R R K E H A
21789 CTCGCTGAGGACATGCTCGCGCGCATGGTACAGGAGTCCCAGAGTTCGCGCGGCTGACGGCGAAGCGGACGTACCCAGCGGGCGTGGCGG
> S L E D M L A A L V Q E S P T V R G L T A N G T Y Q P G V R
21881 TGGAGCAGGACTTCTCGTACATCTCCGACAGCTTCTCGCGGCGCGGCTACTTTCGCGGCGCGGCTGCTGCTGCTTCTCGGACCCACTGCTG
> V E Q D F S Y I S D S F C G P G Y F A A G D S A C F L D P L L
21973 TCCACCGCGCTGACCTTCGCGCTTACAGCGGATGCTGCGCTTCGCGGCTTCACTTCGCGCACCATTCACCGTACGCTACCCAGGAGGAGGG
> S T G V H L A L Y S G M L A S A S I L A T I H G A D V T G E E A
22065 GCGGCGCTTCTACGAGTCCCTTACCGCAACGCTTACCAGCGCTGTTACCCCTGCTGCGCGGCTTACCAGCAGCAGGCGCGCAAGAGGG
> R A F Y E S L Y R N A Y Q R L F T L V A G V Y Q Q Q A G K R
22157 CATACTTCGCGCTTGGCGAGCGCTGGTGCACGACAGCGGCGAACCGAGTACGAGAAGTAGACGGGGCGCGGCTTCGCGGACCTTCGCTC
> A Y F F G L A D A L V H D S G E P E Y E K V D G A R A F A Q L V
22249 CGCGGCTTCGCGGCTGACGAGCGCGGCGGAGGACGACGACGACGCGCGGCGGCGGCGGAGCAGGACCACTTCGCTTCGCGGCA
> A G L A D L D D A E G R H D S T A A A A P A E Q D N S V R Q
22341 GCTCTTCTGCGCGGAGGAGGCGCGCGGATGGCGGACGCGCGACGCGGAGCGCGCGGCTGAGCGAGGCGCGGCGGAGCTGACAGCC
> L F L A A E A E A R R M A D A R T P S A P V S E A P G K L D S
22433 ACGACCTCTTCGACTCGGCAACCGGCTCTACCTGGTACCACCTCCGAGCTGGGAGTCCGCGGGGCAAGCGGCGGACACGAGGCGGCG
> H D L F D S A T G L Y L V T T P R L L G I R R A K P A D T Q A A
22525 GCAGAGCAGTTCGCTGAGGTTCCACCCCTGGTGGCCCGCGGCGGCGGCTCGGGGCTGCTCAACCCCTCCCAACCAACCTCG
> A E Q S A .
22616 GCATCCGTTGCGCGGCTGAGCAGGGGACGCGCCACCGTCCGGCCCGTTCACATGGACAAGGTACCTCTCCCGTGTGAACAGACGACA
22708 GTTGCTCGCCCTCGGACGCTGGCGCGGCGGCGGACCGTTCATGTACGCTCCTCTAGACGGGACGCGAGGCGCGCCAGGCGGCAACCGC
> M S R S L R R D Q Q A A Q A A P A
22798 GTCCCGGCGCAACCGCAGCGCGCGCGCGCGCGGTCGCGGTCAGCAGCAGCAGCGTTCGCGGTCACCCCGTTACCCGAGGCGGA
> S P A N P H A G H A A P V P S R V S T T T V A V T P F T E P
22890 TGCCGCTCCCGCGCGGCTGACCCCGGCTTCCCGCGCGGACGCGGATGACGCTTACGAGATCCCGATCCCGCGCGGCGGAGTGCAGATCTCG
> M P V P P R L T P V S R R D G I D V Y E I P I R P A Q V Q I L
22982 CCCGGCTGCTACGCGCGCTACACCTACCGCGGTTCTTCTGCTCGGCGCCACCATTCGCGCGCGGCGGCGGCGGCTGCGGATCACCTA
> P G L L T P A Y T A G S F V G P T I R A R T G R P V R I T Y
23074 CACCAACGGGCTGACACCCACGCAACCTGACCTGCAGCGGGGACGCTGCGCGGCGACGAGCAGCTGACCCGATGGACCTGATCCCG
> T N G L D T H A N V H L H G G H V P A T S D G H P M D L I P
23166 CGGGCGGCTCGAAGGTTCACGACTACCCGAACCTTCAGCGCGGCGGCGGCTGCTGTAACACGACACCCACGCTACGAGGCGGACCG
> P G G S K V Y D Y P N L Q R G A T L W Y H D H T H A Y E A D H
23258 GTCTACCGCGGATGACGCGGCTCTATCTGATGACGACCGCGGAGCATCTGCGCTGCGCGCGGCGGAGTACGAGCTGCGGATCCG
> V Y R G L H G F Y L I D D P A E H L R L P A G K Y D V P I M
23350 GCTGCGCAACCGCCAGTTTCAGCAGCTCCGCGCGCTGCTCTTCGCGCACCGGACGCGGCTCACCATCTGGCGAACCGCAAGGCCGAGC
> L R N A Q Q D D S S G A L V F G H P D D R V T I L A N G K A Q
23442 CTCACTTCGAGGTGGCGCGCGGACGGTACCGGTTCCGCTGCTCAACGGGGCGCTGAAGCACTCTTCGGCTCAACCTGGGCGGCGAACC

Figure 11F

junction marker

51087 CAGCACAACCCGGCCAGGTAGGAAACAGCCCCGCTCAACCACGTCCTCACGACCATTTCTTGGCAGGCGGCGTATCGGGGGGGGGGGGGGGG
51179 TGGGGATTAAATCAACCCCTTGGCAATCTCTCCGCGAAGGATCAACCCCAATTGGAGGGGTGATCGGGGGGGTGGGGCTGGAACCTCACC
51271 ACCGGCGCTGCAGGCACTCGCCCGGGCTCGGGACGTGGTCCCCACTTTAACTTGACCCCTGCGGCGCTCCGAATGGCGCACCCGCAAGGCGT
51363 CGTAAACACCTTAGGTCTGAATGGTGAGCCGTCACCGCTCATCCACCCGAGCAAGGCGATCTGACGGCGCTACTGCGGATCCCGACCATC
51455 CATGGGCAACAGCGAGGGCGTCTCTCCAGTAGTCCGCCCCCGGAGGGCGGAGGACCAATCGACCCAGGGCCAGCTATTGGCTTCGACCCG
51547 TGA AAAACGCTCTCCAGCGGGCGCTACCTTCACCTTCCTAGCAACATGTATGTTGCATCAGCGTACGCCCCGAGCGCGTGCCTGCGA
> V R A
51637 GTCTGAAACCGATGGTGCCTTGCGAATACGGGCTGACCAGTGGCACCTTAACCTACCTGAAGGCGGCGAGGCTCAGAAGCTGTCACTGGGA
> S P E T D G A L R I R A D Q W H L N Y L K A A E A Q K L S L G
51729 GAAGGGGTTGTAGTAGCGGTCCCGGATACTGGCGTTGATCCACACCCCGACCTTCAGCGCAATCTAATCAAAGGGATTGACATCATTCCCGG
> E G V V V A V P D T G V D P H P D L Q R N L I K G I D I I P G
51821 GGGCAATGGAGATGGCCAGAAAGATCGCAACAGTCAAGGCTACGATGGCTGGACTAATCGACCCACGGACAGGGCCAGAGCGGCGCT
> G N G D G Q K D R N S H G T S M A G L I A A H G Q G Q S G A
51913 TAGGCATAGACCCAGAGCAAGATCATGCCAATCTCTTCCCGCTCGAACAACCTCGGTGATGCAGACGGCTTGGCTGCGGGTATAGAA
> L G I A K I M P I L S S A S N N L G D A D G L A G I E
52005 TTTGCAATCTCGCATGGGGGGGATGTCTCATATGTCTCCAGCGGAGGCGGGCGCCAGCGTTCGACTCATCAAGGCAATCAGAGAGGCGGTGCG
> F A I S H G A D V I N V S S G G A S V R L I K I R E A V A
52097 CGCAGACATTGTAGTTGTCTGCGAGCGCGCGGAACAGTCCCGAAGACATGACAATTGGCTATCCAGCCAGCGAGGAAGGAGTCTGCGCAGTTG
> A D I V V V A A A G N S P E D M T I G Y P A S E E G V V A V
52189 GCGGAATTGATCGACAGGGAGGATGCTTCAGTTCTGCTCGGACCGGAAGTTGACTTAGTTCGACCCGCGAGTCGACATCTACAGCAC
> G G I D R Q G E H A S V S V V G P E V D L V A P A V D I Y S T
52281 AGTTACGACGGGAAGTACTCCAAAGGCACCGGTACGTCCAGTGCCACAGCGATAGTCGAGGGGCGCTGCTCTGGTCCGATCGAAGTTTCC
> S Y D G T S K G T G T S S A T A I V A G A A A L V R S K F P
52373 CGACCTGCCCCGCTCGGAGGTGCTGCTCATCGCTCACTGCCATCGCCATCGACAAAGGGCGCGCAGGGCAGCAGCAGCAGCAGCGGTACGGCG
> D L P A S E V V H R L T A T A I D K G P P G H D Q Y G Y G
52465 TTATCGACCTGGTTGCGCGCTTACGCGAGACGTACCCCGGTGGGCTTTGAGTGGCGCAGCGCGGACGTGCCCGACGTGCTCGGTGCGAC
> V I D L V A A L T A D V P P V G F E S A T A D V P D V P G S T
52557 ACAGCGGCGGTGCGCCAGCGCGGAGGTCAGCATGGGCAAGCGCGGAGTCTGGCCAGCTTGGGAGTATCGTGGCTGCTGCGGG
> T T A V A E P A G E G D D G A T A R G L A T L G V I V A A A G
52649 CGCTTGGGCGGTGGTCTGCTGACGGCGTAGGTTGAGCGACGACCCCGCGCGGATCAGCCGCTGACCGAGCCAGCCCTGACGCCATGTG
> A W L A V R R R L S D D P P P R I S R
52740 GCGACATTGGGGGGTGGGGTGGGTGGGATACCGCTATTGGCGGACATGAAGTCGATCAACAGCGATCAACAGGCAAGTGTGATCGGCGGG
52832 GTCGGTCGACGGCCAGGCGGCTGACGAGGCGGGGGAAGCGGTAGCGTCGGCGACGTGCCGAACCTCGATTTCACTCCGACTCGTGCTCG
> V P N S I S L A L V L A
52924 GTCGGCGAGCCCTGCCCCGTGCAAGCTCTCCACGCGCGCGGATCGAACCAGCGTGTGGTCAAGTGGGGTGCAGAGTCCCAAGTGACCA
> S A S P A R R K L L H A A G I E P D V L V S G V S T
53016 GCGAGGACCCGAGGATCTGTGCTTGGAGCTGACCCCGTGAAGGCGCAGGCGGTGCTGCGCGGCTGCGCCCGTCCGCGGACGAGCGGACG
> S E R A E D L C L E L A R L K A Q A V V G R L R P S A D E R T
53108 CTGGTCTCGGCTGCGACTCGGTGCTGCTGCGCTTCGACCGGAGTCTCGGCAAGCGCGGACGAGCGGAGCTACCCGCGCTTGGGAGCG
> L V L G C D S V L A F D R E I L G K P A D E A D A T R R W E R
53200 GATCGGGGGCGCAGCGGGGTGCTACACACCGGGCAGTCCGTGATCGACGTATCCACGAGACGCGCGCGGAGGCGGTGCGCTCGACCCCG
> M R G R S G V L H T G H C L I D V I H E T R A E A V A C T T
53292 TGCCTTTCTGCTGACATCAGCAGCAGGAGATTGCGCGTACGTGCGCAGCGGCGAACCCTGCGGCTGCGCGCGCTTACCATTGACCGGA
> V R F A D I S D E E I A A Y V A T G E P L A V A G A F T I D G
53384 ATGGGCGGGCGCTTCTGGAGGGTGTGACGCGGACCCGCGCAGGTGGTGGCTCTCCCTACCGTGTGCGCGCGCTTCTCGGCGAGCT
> M G G A F L E G V D G D P G T V V S L S L P L L R R L L G E L
53476 GGACCTGCGGATCATGACCTGTGGACGAAGGTGCGCGCGGCGCGCAGGCGGTGAGGCGGTGGGTACGCTCAGCGATGACGAGGAGT
> M T T K
> D L R I A I D L W T K V A P G G Q A V E A V G T V Q P
53567 CCCTGCGCTGACCCCGGAACCTGCAATGCGTACGTGGTGGCCACGGATCGGACCCGAGGAGTGTGATCGGAGGAGACCCCTC
> S L P L T P E L H A Y V V A H G S D P D E V M R D L I E E T L
53659 GCGCGCTGCCCCCGGAGGCGAGGATGCAAGTGGCCCGGAGCAAGCGCGTTCCTGACGTTCCTCACCCGGTTGATCGGGGCGCGGCGGCG
> A A L P A E A R M Q V A P E Q A A F L T F L T R L I G A R R A
53751 GGTGGAGGTGGGACCTTCACCGGCTGTCTCCCTGGCGATCGCGCGCGGCTGCGCGAGGCGGCGGCTTACCTGCTTCGACATCTCGG
> V E V G T F T G L S S L A I A R G L A E G G R L T C F D I S
53843 AGGATACACGCGTTCGCGCGGTACTGGCGCGGGCGGGGTGGCGGACGAGATCGACCTGCGGATCGGGCGGCGGGGACACCGTG
> E E Y T G V A R R Y W A R A G V A D Q I D L R I G P A G D T L
53935 CGCGGTTGCGGTACGAACGCGACCTGGACTTCGCGTTCATCGACCGGACAAAGGTGCGCTACCCGCTTACTGGCGGAGTGGTGCCCG
> R G L P Y E R H L D F A F I D A D K V G Y P V Y W A E L V P R
54027 CATGCTCCCGGGCGGGTTCATCGCGGTGACAAACAGTTCGCGCGGGCGGGTGTCTGCGCCCGGCTGACGCGGACGACCGGGCATCGCG
> M L P G G V I A V D N T L R G G R V L A P R D A D R D A I A
54119 CGTTCAACGACGAGGTGATGGCCGACGTCCGGGTGAGCGCGGTGCTGCTGCGCATCGCGGACGCGGCTGACCTGGCCCGGTGCGCTGACG
> A F N D E V M A D V R V E P V L L P I A D G L T L A R V R
54210 GCGCGGACCGGAGATCGTGGCAGGTGCGCGCGCGCGCTGCTGTGCTGCGCTACCGGGTTCAGCGAGGGGTGAGGCGGGGCGGGGTGT
54302 AGGAAGGGGCGCTTCTATACCGAATGCGTTAACAAGGGGCGCGCTTACACCTCAACCTCAGCGCACGCTGCGGCGAAGTTCGCGGCG
< R V S R A Q R A
54393 GCGGCGGACGCGGACGCGCGGCAAGCAGCGGATGATGGTCAGGCGCTGCGAGACCTTGTCTGTTGCGGAGGTGCGCGGCGAAGAGGGCGCG
< A W A V G V A A L V A I I T L G Q W V K D N G L D G A F L A R
54485 GGTGCGCTCACGGCCAGGAGAACGGGTTCACTTCGCGATGCGCTGGAGCCAGCGCGGGGCGAAGGTGAGCGGCAGCAGGATGCGCGGAG
< T G D V A W S F P N W E A I R Q L W G P A F T L P L L I G S L
54577 GCAGCAGCAGCGGTGCGGCGAGGTTTCATCACCGGGCGAGCGGTCTCTACTCTTGACCTTGAGCGCGACGCGGTACGAGACGCGCGAG
< L L V P Q A V T N M V P A L A D E S K V K L A V G S V A S
54669 GTCATCAGCGCATCAGGCGAGCATCAGGTACGCCAGCAGCGGTGCGCGATGAACACGCGCAGCTCGAACAGGAGCGCGAGCGAGGTTGAT
< T M L A I L A L M L L L D L D G I F V R L L E F L L T I
54761 GATACGGCTTGGGCGAGCAGCAGCAGCTGCGCGAGGCGCGCGGCGGAGCAGCGCGAGCGCGGCTGACCGGGGTGACCGGGGACCGGTT
< I V A Q A L L S V V D R L A R G L L A L R S V P T V R S R E
54853 CGATGACCGCGCGCAGCTCGGCGATGCGCGGAGGCGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
< I V G A R L E A I L G F G Q F L G G F I A L L V L L G P V F
54945 ATCTTGTACGCTCGGCTGGGTGCGGCGGTTCAGCGCGGGCTTGAGCAGCGGGGCGAAGAGGAGCAGGTACATCACCGGCTGGAAGACGCC
< I K Y A E A Q T P A N L A P K L L P A F L L L M V P Q F V G
55037 GACGAAGACCGAGCGGATGCGGAGCAGGATTGCATCTGCGCTGGGCGACGAGCGAGGTGTCGCGGGCGAAGTTCATGATCGGACT
< V F V W V P N R L L L Q M Q R Q A V L W T D R A F K M

Figure 11N

55127 CCGGGTGGTCAAGACTCGCGCAGCGAGCGCGCGTCTTGGTGAGGAAGACGTGTCGAGGCTGGGGCGGTGCAGCTCGATCGAGCTGAGCC
< . S E R L S R G T K T L F V D D L S P R H L E I S S L R
55218 TGAGGCCGACTGGTTCGAGCCGGCGCAGGACCTGCGGGATGGCGGTGGCCCCCTCGTGCAGGTCAGGCGCAGGCCCGCGCTCGACGGT
< L G S Q D L R R L V Q P I A T A G E D V T L R L G G G D V T
55310 TCCAGCTTGGTACGTACGGTTCGAGTCAGTTCGGGTGCGGCGCGCTCCAGCCCGACGAGCAGCACTCGCCGGA
< E L K T V Y P E T D L L Q A A Q P T A A A D L G V L L V E G S
55402 GATCTCCCGCTTCAGCCCGCGCGGTACCTCGGCGAGCACCCTGCGCGTGGTCCATGATCGCGATCCGGTCGAGAGCGCGTCCGGCTCGT
< I E R K L G G P T G E A V V E G H D M I A I R D C L A D A E D
55494 CCAGGTAGTGGTGGTGAAGACGGTCATCCCTCGGCGCGCAGCGCAGGATCTCGTCCACATGTGGGCGGACTCTCGGGTTCGAGG
< L Y H T T I F V T M G E A R L R R I E D W M H A R S Q P D L
55586 CCGCTGGTTCGCTCCAGGAACAATGCGGGGTGCTGGATGATGCGGAGAGCGATCTCGACGCGCGCGCTGGCCCGCGGAGTGGT
< G S T P E D L F V I R P D H I I G L A I E V R R R Q G G S Y T
55678 CTTGCACTTACGGTCGGCTACTCGGTGAGCTGGAAGCGCGGCGAGTGGCGCTCGGCGCGCGGAGGGCGTGGCCCTTGGCCGATGCGGTACA
< K C K R D A Y E T L Q F A A L A R E A R R L A D A K D G I G Y M
55770 TCCGGCGTGCAGGACCACTTCTCGCGGGCGGTGGAGTCTGCGCAGGTCTGCGCCCTGGGCGACATAGCCGATCCGCGCAGCAGCACTCG
< R A H L V L E E R A T S D D W T S G G Q A V Y G I R R V E
55862 GCGCGGTTCGCGAGGTCGGCCCGCGGTGGTGGCTGGCCCGCGCGCGGCGGTGATGAGGTGGCGAGCATCCGAGGGTGGTGGTCTT
< A P N R L L D A G A I T A Q G G D P T I L T A L M R L T T T K
55954 CCGGCGCGCTGGGGCGGAGGAACCGAAGATCTCCCTCGGCGAGCTCCAGGTGACGCGCGCAGCGCTGACCGTCTGTGTCTGTC
< G A G N P G L F G F I E G A V D L D V G R V A D V T K H Q R
56046 GACCGCGCGGGAGCGAAACGACTTCGCGAGCCCTCTGGTCTGGATCATCTTCTGCTCTGGTCTCTTAGCCGAGCGGGCGCGCCCTC
< G A R S R R C F S K R L G R T Q I M
56136 TCTCCGGGACGCGACACCGGTCGCGCCCGAACGTCGCGCGGAGGCTAACCGGATATAACTTCTTAGTCAACTTTGATTAATGGCGA
< . R S I V E R T L K S . H R
56227 CCGTCGCGCCCCCTCCCCACGTTCCAGCCGCTCTGACTGGCCAACTTCGGCGAGATACGGCAGCGCGGCTCGATCCGGTCCGCGACCCG
< G D A G E G V N W G D Q S A L G E P L Y P V G A E I R D A V R
56319 CTCACACAGGCCACTCGACCTCTCCCCGGGAATCCACAGCTCGTATCCAGCTCAGCGCCAGCGGCTTGGAGTCCGGATCCAGGAGG
< E C W A V E V E G R A I W L E Y M W S V G V P K S S S
56411 ACTCCATCGAGGACGCGATGGTTTCGACACTGGCCCGCAGCAGCTGCCCCGACTGCGCAGCGCGGCCACCGCTCGGCGTGGGCGAGCGCC
< E M S A R M T E V S A R L V Q G R S R L A A V A E A P R P L A
56503 GGCAGGAACGCGAAGCGCGCCACGAGATCGCTCGTATGATGTCGCCACACAGCGCGCAGCAGCTCTCGAATCTCGTACCGCC
< P L F A F A A V F P D S T Q H N G W W L G R L L T E F E D V G
56595 CTTCCGGGTGATCTCGTACGTCTCGCGCGCGCGCGCGCGCGCTGCTCGGTGGCGACCTCGCGGAGCAGCGCCCTCTCGCGGAGCTTGC
< K P T I E Y T T R A R R A G V Q T C T A V E R L L G E E K R
56687 GCAGCGGTGGTAGATCGAGCCGGGTGCACGTGGGCCACTTGTGCGCACCCCACTGAGCAGCTCGCGCGGAGCTCGTAGCCGTGCACC
< L A H Y I S G P Q V N A W K D A G W S L L E R R V D Y G H V
56779 GGCTGCATCCACTTGACAGGCGAGAATCATCATGAGTGGCAGACACCGGAAAAGCGTATTAGACAAGTTTGACTATCCAAAGCATCTG
< P Q M W K V L G L I M M
56870 GGCAGTGCCTCATCCACACTGAGCGATCGTTAGGGCCACGACGCGCGCGGATAACTCCCCGTGAGTAACATCCCGGAGGAGCCACGAG
56961 GTGCGCAAGGTACTCATCGCAACCGAGGCGAGATCGCGCTCGCGCTCATCCGCGCTCGCGCGAGCGCGCGCTCGCGCTCT
> V R K V L I A N R G E I A V R V I R A C R D A G L G S V A V
57052 ACGGGAGCTCCGACCGGGACGCCCTGCACGCGACCCCTGGCCGAGAGGCGTACGCCCTGGGCGGCGACACCGCGCGGAGACGTACCTGCGG
> Y A D S D R D A L H A T L A D E A Y A L G G D T A R S E T Y L R
57144 ATCGACAAGCTGATCGCCGTGCGCGCACAGGCGGGCGGCGCGCTCCACCCGGGTAGCGCTTCTCGCGGAGAACCGCGACTTCGCCCCA
> I D K L I A V A A Q A G A D A V A H P C Y G F L A E N A D F A Q
57236 GGCCGCTCGACGCGCGGCTTACCTGGAATCGGCGCGACCCACAGGCGATCCGCGACCTGGGCGACAAGGTACACCGCGCGCACATCGCCC
> A V L D A G L T W I C P T P Q A I R D L G D K V T A R H I A
57328 AGCGGGCGCGCGCGCCCTGGTTCGCGGTACTCGGACCGCGGTGGCGAGCGCGGACGAGGTGATCGCATTCGCGGTGACACCGCGCTGCGG
> Q R A G A P G T P G T S D P V G S P D E V I A F A V D H G L P
57420 GTGCGCATCAAGCGCGCTTCGCGGGCGCGCGGCGCGGCTCAAGGTGGCGCGCACGATGGAGGAGATCCCGCACCTGTTTCGAGTCGCGCAC
> V A I K A A F G G G G R G L K V A R T M E E I P H L F E S A T
57512 CCGGAGGCGGTTCGCGCGGTTCGCGCGGGCGAGTGTTCGTCGAGCGGTACCTCGACACCGCGCGCACGTCGAGGCGCGAGGTCTCGCGG
> R E A V A A F G R G E C F V E R Y L D Q P R H V E A Q V L A
57604 ACCAGCACGGCAACGTGATCGTCTCGGACCCCGGACTGCTCGTCAACGCGCGCACGAACTCGTCGAGGAGGCGCGCGCGCTTC
> D Q H G N V I V G T R D C S L Q R R H Q K L V E E A P A P F
57696 CTCACCGCGCGCGCGCGGAGATCGACGACGCGCAAGGCAATCTCGCGGAGGCGCGCTACACGCGCGCGCGCACCGTGGAGTACCT
> L T D A Q R R Q I H D S A K A I C R E A G Y H G A G Y L
57788 GGTGGGACGCGACGCGACGATCTCTTCTTGGAGTCAACACCCGCTGCGAGGTGAGCAGCCGCTACCGAGGAACCGCGCGCATCGACC
> V G T D G T I S F L E A V N T R L Q V E H P V T E E T A G I D
57880 TCGTCCGTAGCATTCGCGATCGCGACGCGAGAACTGCGGTGGCGGAGGATCCGACCCCGCGCGGCACTCCATCGAGTTCGCGATC
> L V R E Q F R I A D G E K L R L A E D P T P R G H S I E F R I
57972 AACGGCGAGGATCCGGCGCGCAACTTCTGCGCGCGCGCGCGCGCTACCGCGCTGCGGCTGCGCCACCGCGCGCGGTGTCGCGGTGGACAC
> N G E D P G R N F L P A P G T V T A L R L P T G P G V R V D T
58064 CCGCATCTCCGCGCGCGACGTGATCGCGCGCAACTTCGACTCCTGCTGGCAAGGTGATCATCACGGGCGAGACCGCACCGAGGCGCTGG
> G I S A G G D V I G G N F D S L L A K V I I T G E T R T E A L
58156 AGCGGGCGCGCGCGCTGGACGAGATGGTCTGAGGGAATGGCCACGGCGCTGCGGTTCCACCGCTGGTGGTACGCGACCCCGCGTTC
> E R A R R A L D E M V V E G M A T A L P F H R L V V R D P A F
58248 ACCGCGCGCGCTTACCGTGCACACCCGGTGGATCGAGACGGAGTTCGACAACACCGTCTGCGGTTACCGCGCGCGCGCGCGCGCGGCGG
> T A A P F T V H T R W I E T E F D N T V L P F T A A A G P A E
58340 GGGCCGCGCGGAGCGGAGACGCTCGTGGTTCGAGGTGGCGCGCAAGCGGTGAGGTGACCTCCCGCGCGCGCTCGCGCGCGGTACGCGCG
> G P A E R E T V V V G G K R L E V T L P A G L A G T A
58432 CCGGGCGCGCGCGGAAGCG
> A G P A A R K P A R R G G G A K A G A A V G G D A L T S P M Q
58524 GGCACGATCGTAAGATCGCGGTGCGCGACGGGACACCGTCCGCAAGGGCGACCTGGTCTGCTGGAGCGATGAAGATGAGCAGCC
> G T I V K I A V A D G D T V A K G D L V V L E A M K M E Q P
58616 GCTGCACGCGACAAGCGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGT
> L H A H K A G T V G G G C T A A G G A V L A A G A P I C T I T
58708 GAGGTCAAGGAGGGGCGCCCTGTAAACGATTTCGGTATAGGAAGGGCCCTTCTTAACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
> .
58800 TACGCGTACCGCGCGGGGTGTTTTCCGCGACACCGCGAGCGGTGAGGACCGGGCGCGGAATGATGGCCAGGTGCGGTTCCTACATGGC
> > V R F L H G
58891 GCGGTTCGCGCGCACGACCTGACCTACAACGACGTCTTCATGGCGCGCAACCGCTCCGAGGTGCGGTTCGCGGTGAGACGTGACCTGGCCAC

Figure 110

> A V P A H D L T Y N D V F M A P N R S E V G S R L D V D L A T
58983 CTCCGACGGCACGGGCACACCATCCCGCTGGTGGTGCGCAACATGACGGCGGTGGCGGGCGGCGATGGCCGAGACTGTCCGCCGGCGGG
> S D G T G T T I P L V A N M T A V A G R M A E T V A R R
59075 GCGCACTCGCGGTGATCCCGCAGGACATCCCGATCGAGGTGGTGGCCAACTGGTTCGCTGGGTCAAGCAACGGCACCTGGTGCACGACACG
> G A L A V I P Q D I P I E V V A N V V A W V K Q R H L V H D T
59167 GCGATCACGCTCGGCCCCACCGACCTGCGGATGCCATCTGCTGCCGAAACGGTGCATGCGCGGTGGTGGTGGTTCGACGAGGC
> A I T L G P T D T C V G D A I H L L P K R S H G A V V V V D E A
59259 CGGTGCGGCGCTGGGCGTGGTGAAGGACGGGACCGCTCGGGGTGAGCCGCTTCGCCCACTCGGCCACGTGATGTCGACCGGATGGCACA
> G R P L G V V T E A D T V G V D R F A Q L R H V M S T E L H
59351 CGGTGCGGCGGACGCGGACCCGCGTACCGGATTTCGACCGGCTCTCGGCGGGCGGCGGCGGCTCGCGCCGGTGGTGGACGGCGACGGCCGG
> T V P A D A D P R T G F D R L S A G R R R L A P V V D G D G R
59443 CTCGTGCGGCTGTTGACCGCAAGGGCGCGCTCGCGCGACCCCTCTACACCCCGCGGTGGACACCGGGGCGCGGTGCGGATCGCGGCGGG
> L V G V L T R K G A L R A T L Y T P A V D D R G R L R I A A A
59535 CGTCGCGCATCAACCGCGACGCTACCGGCAAGCCGCGCGCTGCTGGAGCGCGGGCTCGACGCCCTGGTGGTGGACACCGCGACGGCCACC
> V G I N G D V T T G K A A A L L E A G V D A L V V D T A H G H
59627 AGGCGCGGATGCTGCGCGCGCTGCGGGCGGTGCGCAAGCTTACCCGGGCGTTCGGGTGCGCGCGGCAACGTGGTACCCGCCGATGGGGTA
> Q A R M V A A L R A V R K L H P G V P V A A G N V V T A D G V
59719 CGCGACCTCGTCGAGGCGCGCGACATCGTGAAGGTGGGCGCTCGGTTCGGGCGCGATGTGCACACCCGGATGATGACCGGGGTGGGGCG
> R D L V E A G A D I V K V G V G P G A M C T T R M M T G V G R
59811 TCCGAGTTCCTCCGCGGTGCTGGAATCGCGCGCGCGCGGACCTCGGCGGACGTCGCGCGGCGGTGAGGCGGCGGTGACGCGGCGCGCG
> P Q F S A V L D C A A A A R D L G R H V W A D G G V R H P R
59903 ACGTGGCGCTGGCGCTCGCGCGCGCGCTCGAACCTGATGATCGTTCCTGGTTCGCGCGGACGTCAGTACGATCCCGGGTGACCTGACACG
> D V A L A L A A G A G A S N V M I G S W F A G T Y E S P G D L Y T
59995 GACGCGGACGCGCGGAGGTACAAGGAGAGCTTCGGGATGGCTCGTTCGCGGCGGCTCAGCGCGGTACGCGCGGAGGACAGCGGCTTCGACCG
> D A D A G R R Y K E S F G M A S S R A V S A R T A E D S A F D R
60087 GGCCGCAAGGGGATCTTCAGGAGGGCATCTCCTCGGCCGGATGTACTCAACCCGATCGCCCGGCGGTGAGGACCTGATCGACGAGA
> A R K G I F E E G I S S A R M Y L D P D T G R P G V E D L I D E
60179 TCATCTCCGGGTACGACGCGCGTACGCTACGCGCGCGCGCGAGCTGGCGGAGTTCGCGGAGCGGGCGCTGGTGGGGTGCAGAGCAGC
> I S G V R S A C T Y A G A R S L A E F A E R A L V G V Q S T
60271 GCGGCTACACCGAGGGATGCCCTACCGACGAGTGTGGTACCCCGCGCGCGGCGGTGAGAAGGGTTCCTCTCTACCGGAGCGCTCAA
> A G Y T C G M P L P T S W
60362 CAAGGGGCTTCCTTCGTGCGCGGTGGTATCGCGGTACCGACTGCGGCACGCGCGCGCGCACACTGAGCGCGCGCGCGTTCGAGGGGCC
60454 ACCGAACGGCGCGCGGGTCACTCGAAGAGGCGACGGATGACGCTCGCGGCGCGCGCTCGGGTTCGGGCGCGGTGCGCGGCGGAGCGGCC
< . D F L R C R I V T R A A A E P D P G T G P P L A
60545 CCGGCCAGCCAGAGTGTACGAAGCGGTGACGATCGACGAGCGCGCGCGCGCGCGCACACTGAGCGCGCGCGCGTTCGAGGGGCC
< G A L W L T V F G H V I S W A A L A D A E Q D P D T E R R P L
60637 GGCGGCCACCCCGCGCGCGCGCGCGCGCGATCACGGCGGCTGTCACTTCGGGTGCTGCGACGGGTAGAGCTTCGCGCGGGAACA
< A A V G A R L A A G A R D R A A T V E P D D R R Y L E P R F M
60729 TCACCTCGAAGTGGGCGCGGTGCGACGCGGCAACCGGACGCGCGCGCGGTGCGAGAGGTGCGCGCGCTCGCACAGCGCGCGCGCG
< V E F H A R H D V A F R V Y A V G A D L L D G A E C L A G A
60821 AGCAGGTGCAATCCCTCGACGGCGAGCGCGGTGAGCAGCGCGCGCTTGTCGCGGAAGTGGTGGCGGGGGCGCGTGCAGACCCCGGCGCG
< L L D F G E V A L A T L L G A K D G F H H A P A A H S V G A R
60913 GCGGCGCAGGTGCGCAGGCTCAGGCGCGCGCGCGCGCTGCGGTGATCGCGTGCAGCGCGCGCGGAGCAGGCGCGCAGCGAGTCACTCT
< R A L D R L S L A A P G A D T I A D V A A A L L A R R L D G H
61005 GATGGTAGCCACGCGGTTCGGGTCACTCGCGGAGCCTAATTGTCTATTGACAAGATAGCCAGGCGGAAGCAATCTAGGCAATGACAAGTTG
< H Y G R P G T M
61095 CCTTCGACCGAGGAGAACCCCGATCGCGCGCGCTGATCGCTCTCATCGCGGCTTCGCGCGCTGGCCGCGCTACGCGGCTACTCAACGTGCGAG
61187 CCTGCGCGGCTGCGCACCCCGCGCTCGGGTTCGGGCTCGCGGCTGCTTCGCGCTACCGGATCGCCACTTCACTCGACGCGCGCGAC
61279 CCGTGGCGGCTGCGCGCGCGCGCGCGCGCGCGCTGCTGGTACCGTCAACCGCGCTGCTGGAGTTGGCGCGCGCGGTTCGCGCTGCT
> M V P P R L P H P G L L V T V T G L L E L A G A V A L L
61371 CGTCCCCGCGACGCGGTGGGCGAGCGCGCGGTGGGCTGCTGCTGCTGCGATGTTCCCGGCAACGCTCGCGCGCGCGCGCGCG
> V P G T A R W A A A G L G L L L A M F P A N A S A A R R G
61463 TGACCTTGCGCGCGCGCGCGGTGACCGCGCTGCTCCCCGCGCGCTGCTCAGGTGATCTTCTCACCGCGCGCGCGGATTCGTTGGG
> L T L A G R P V T P L P R A L L Q V I F L T A A A A I S F G
61555 CCCTGACTATCAGGGAGCTAACATGACCCGATGGAGCTGATAGGGCTGCGAGACGTGCCACTCGGCGGCTGCTGGTGACCGCGCGCA
> P . > V P L G R L L V T A G H
61645 CGTCTCGGCCAACGGTGGAAACCGCTACCTCGCGGAGGAGCAGCGCTCACCCAGGCGGATGGTACCCCTGATACCTGGCGCGGACG
> V V G G R W N R Y L A E E H G L T Q A G M V T L M T L A R H
61737 GCGAGTGCAGCGCGGCGGTGCGCGAGGCGGTCTTCACTCGCGCGCGCGCGCGCTAACCGGATCGTCGACACACTGGAGCGCGCGGCTC
> G E L P H R A V A E A C F I R P A T L T G I V D T L E R D G L
61829 GTCGAGCGCAACCGCAGCAGCTGACCGCGCGAGCGTGGCGCTGCTGACCCCGCGCGTTCGGAACCGGTGCGCGCGCTCACCACGT
> V E R Q R D D V D R R S V R L V L T P A G R E R V A A L T N V
61921 CATCGAGTCCGAGCAGGACGATGACCTCGGTGCGAGCGCGACCCCGGCAAGCGCGCGTTCGCGGAGTTCCTGCTCGAGGTTCATCGGCGAGT
> M Q S G R P M T S V D A D P A K A A V I R Q F L L E V I G S
62013 GAGAGGAACCTCGGCTGACGCGCTCGACGCGAGGCGGAGGCTCGCGCATGTCTCGGCTGCTCGCGCGCGCGCGCGCGCGCGCGCG
> G E E P R V T A L D A R P E A P A C
62105 GACCGCTGCGCGCGGTGATGGCGTTCAGTTCGTCGCGGACGATGGCTTCGCTACCTGCGGAGCTCAACCGCGGACATCATCGACGAGG
> M A L Q F V G T M A S L Y L L P S L N A D I I D Q G
62196 TGTGGCCCGGGCGACACCGGCTACATCATCGGTACGGCGCGGTGGATGCTGCTGGTACGCTGGTGCAGATGCCTGCTCCACCGCGCGCG
> V A R G D T G Y I M R T G G W M L V L S L Q I A C S T A A
62288 TCTTCTCGCGCGCGCTCGCGATGGGCTTCGCGCGGAGCTACGCGCGGAGTCTTCGCGGACGTCAACCGGTTCCTCGCGCGCGAGGTG
> V F L G A R S A M G F G R D V R A E V F A H V N R F S A R E V
62380 ACCCGCTTCGGCGCACCTCGCTGATCACCGCGCAACACCAACGAGCTGCAACAGGTGCAGATGCTCGTCTGATGAGCTGCACCATGCTGGT
> T R F G A P S L I T R N T N D V Q Q V Q M L V L M S C T M L V
62472 CGCGCGCGGATCATGAGCGTTCGCGGGGTGTTATGCGACTGCGGAGGAGCGTGGCGGTGCTCTGGTGTGCTGGTTCAGCGTGGCGCGCG
> A A P I M S V G G F M A L R E D V V G L S W L M L V S V P A
62564 TGGCGATCGCCCTGATGCTGATCATCGGCGGATGGTGGCCGGGTTCGCGCTGATGACAGCCGCGATCGACGCGGTCAACCGCGTCTGCGCG
> L A I A L M L I I R R M V P G F R L M Q T R I D A V R N V R
62656 GAGCAGTACCGGATCCGGTGGTTCGCGCGTTCGCGCGGAGCTACGAGACGGCGCGCTTCGCGCGCGCGGCAACGCGGACCTCACCGC
> E Q I T G I R V V R A F V R E P Y E T A R F G R A N A D L T
62748 GACCGCGCTGCGCACCGGTGCTGATGGCGCTGATCTTCCCGTGGTGAAGCTGGTGTCAACGCTCCAGCGTTCGCGGTGCTGGTGGTTC
> T A L R T G R L M A L I F P V V T L V L N V S S V A V L W F

Figure 11P

83237 AGCCCCGGGCGCAGGGGCGCGTGCCTCAGACGGCCCCGGGCGCTCAGGCGCGCTTGCCGAGGGCGGCGTCCAGCCGCTTGCCGAGGGCGGT
 < . A R K G L A A D V A K G L A T
 83328 GACGACCAGGGCCACCGAGGGGCGGACCACCGAGTCTGCTCCAGGTGACGGTGCCGGAGAGCCGGCCCCGCGGATCTCCTCCAGCCGAC
 < V V L A V S P R V V S D D L S V T G S F G A G G A I E L R R
 83420 GCCGGGCGTCCGACGGCGCGCTGCGCGGTACGGCCGAGCGCGCGGTCCATCCGGAGCACCAGGAGTGGCGAGCGCGCGGCTCGCTCG
 < R A D V A A D G T L G L A P E M R L V A V L T A L A A T A E
 83512 TCGGGGATCTGCCCCGCGGCCAGCGCTCGGCCAGCGCGCGCGGAGTCTCTCCACGGTGGCGTCCGCGTCCGGTAGCGGTACCGTG
 < D P I Q G G A L A E A L R R R S D E E V T A D A T P Y R H V H
 83604 GATGAAGCCAGCTCGGTCTCTCGACGTTCGCGCACCAGCGCGCGGCGGAGGTCGCGCAGGATCCGGTTCGCGCAGGCGGTGGCGCAGCC
 < I F G L E T E E V D R V V G R A V L D G L I R D G L R L R
 83696 GCTGCACCCAGGAGCGGGGTGTGCGGCGTGTGCGCGCCATCCGCGCCAGGACCTCGTCCAGGATCGGTTTCGCGGTTGGCGCGGGGTCC
 < Q V W S S P T H P T D A A M R G L V E D L I P E G T P A P D
 83788 GTGACCACCGATTCCATCGACGTACGCGACCGCGCGGCGAGGCGGATCGATCAGGACGGCGCGCCATCCCGAGGTTCGAGGCTGAT
 < T V V L N G D V Y A V R G A L A L E I L V A A A M G L D L S I
 83880 CCGCGGATGGTTCGCTTCCCGGATTCTGCTGCTGACGCGAGGAGCAGCAATTCCTCGGCGAGCGCAACACAGTCATGGCCGGAGACGG
 < R P M
 83970 TAGCGCTGAGCGCACCCCGTGCGCCCCCAACTCGCCACGAGCGCACTCGCCCGGTGAGAGGGGAACCCCGCTATACCGCAGGCGTTAACA
 84062 GGGGGCCCTTCTCTGCGATCAGAAGCGGGGATACCGCGGAAGTGGCGGTGCGCGCGCTCGCCGAGGCGCGGACAGTGAACATCCGGTCC
 < . F R P M G G F R G D G A D G L G P V I F M R D
 84153 TTGAGGCTCTCGTTCGATCGCGGCGGTGACGAGGCGCAGCGCGAGGCGGACTGCTCCAGCGGGCGATGCCGACGGCGCGCGGAGCAGCA
 < N L S E D I A A T V L R L P L G S Q E L R A I G V P A A L V C
 84245 GAGCAGGTGATGTCGTTGACGCCCGCTCGGCCAGCAGCGCGGAGTCTCCAGGAGAGCGCGCGGTGGCCAGGATCGGGTTCGAGGACCA
 < L V T I D T C G R E A L L R C C H E L S G G T A L M P D L V L
 84337 GCACGGGCGAGGCGCGGAGTTCGCGCGGCGAGTCCATGTACGCGCGGCTCGTACGTCTCTCGTTCGCGGCGAGGCGGAGCGAAGCC
 < V P L G A L D R P L S E M Y A R P E Y T E E D R A L G V F G
 84429 ATGGACGACTCGGGGAGCAGGGCGAGCGCGCGCTCGGCCATCGCGAGACCGCGCGCAGCAGGGTACCAGCAGCGCGGGTGGCCAGCGG
 < M S S E P L L A L A A D A M G L G A R L V P V L P P N A L R
 84521 GGTGCCCTCGGCTCGGTGACGGGGTCTGACCGGGTACTTCTGACGGGGAAGGAGCGGCGGCTCGTACACGAGCATGGTGGTGGT
 < T G E A D T V P T Q V P Y K E V P F S R A A E Y V L M T T L E
 84613 CGTGCAGCGCGGCGCGAAGTTGGAGGAGTTCGGTTCGCGCGTCCCGCATGGCGGTTCAGCCGCGACTGGGCGAGCGGATGGTCAATGACGTGT
 < H L A S A R F N S D R M
 84705 ACGTCCACGATCGCCCAACCTACCGAACCGCGGGCGCGAGACAGGAGCGACGCCGCTCACCCGCGCGGTGCGGCAACCCCTGGACGGT
 84797 TGGCGTTCCGAAGGTAGGGCAGGTGCCCAAGGTTCGGGCGCGACCTCGCGCGTGTATCAAGATCACGAGGCGTCCGGTTCGCTAGACTTCGGG
 84889 CATGACGGCGACAGCGAGCTCGGCGCGGTTCGGACCTTCGAGCTGGGACGATCCGAGACCGCTTTCGCGAAGCTTCGACCGGCGGTGCG
 > M T A T A T S A R S D L S E L G R S E T A L R N F L H G L P
 84980 GGGCTGGACAGGTTCGGCGCGGAGCAGCGGGCGCGCGCGCTCGGTCACCCGCTCGATCAAGACACCGGCAAGGCGCGCGGATCGACCTGGC
 > G V D Q V G A E Q R A A Q L G T R S I K T T A K A R A I D L A
 85072 GATCCGATGGTTCGACCTGACCACTGAGGGGGCGACACCCCGCGCAAGGTTCGCGGCGTTCGCGGCGCAAGCACTGCGCGCGCGCGG
 > I R M V D L T T L E G A D T P G K V R A L A A K A R P D P
 85164 CCGACCCGCTGCGCGCAGTTCGGCGAGTTCGCTTACCCGCGATGGTTCGCTACGTTGGCGGAGGTGCTGCGCGGATCCGCGGGTCC
 > A D P S C P H V G A V C V Y P A M V P Y V A E V L R G S A G S
 85256 GGGCGGCTCCGCGGACCGGACCGGCAACCGCGCGCGGCGCGGCTGGTTCACCTGGCCAGCGTGGCCACCGGTTTCGCTCCGCGGCA
 > G R P S G G P D G N A P A G P G V V H L A S V A T A F P S G Q
 85348 GGCACCCCTGGAGGTCAAGCTCGCGGACACCGCGCGGAGTGGCGGCTGGCGCGGACGAGATCGACATGGTATCAACCGGGCGCGTTC
 > A P L E V K L A D T R A A V A A G A D E I D M V I N R G A F
 85440 TGGCGGCGCTACCGGAGGTCTACGACGAGATCGTGGCCACCAACAGGCGTTCGCGGACCGCCACCTCAAGGTGATCCTGAAACCGGC
 > L A G R Y R E V Y D E I V A T K Q A C G D A H L K V I L E T G
 85532 GAGCTGGCCACGTACGACACAGTTCGCGCGGCGGCTGCGCGGCGGCGGCTTCATCAAGACCTCGACCGGCAAGGTTC
 > E L A T Y D N V R R A S W L A M L A G G D F I K T S T G K V P
 85624 CGTTCGCGGACCTCCCGGTGACGTGGTGTGAGGCGGCTTCGCGACTTCGCGCGCGGCGGCGGCGGAGGTTCGCGGTGAAGCCGG
 > V A T P L V T L V M L E A V R D F R A A T G F R A A T G V K P
 85716 CCGGCGCATCAAGAACACCAAGGACGCGATCAAGTACCTGGTTATGGTCAACGAGACCGTTCGCGCGGAGTGGCTGGACCGGAGTGGTTC
 > A G G I K N T K D A I K Y L V M V N E T V G P D W L D P D W F
 85808 CGGTTTCGCGGCTGCTCAAGCGCTGCTCAAGCAGCTCATGACGCGACCAAGCTGACGACCGGCGTCTACTCCGTTCCGACTACTTCAACCT
 > R F G A S S L L N D L L M Q R T K L T T G V Y S G P D Y T C T L
 85900 GGACTGAGCGTACTTTCGAATACGCGCGCGCGCGCGCTCCGCTCGGTGGTGGACCTCAAGCCCTCGTACGGGCTGTTCTGTCGACGG
 > D . > V I F E Y A P A P E S R S V V D L K P S Y G L F V D G
 85989 GGAGTTCTGTCGACCCGCGGACGGCGGCGGCTCAAGTTCGCTCAACCCGCGCTCGAGGAGGTGCTCGCGGAGATCGCGGAGGCGGCGG
 > E F V D P A D G G G F K S V N P A S E E V L A E I A E A G S
 86081 CCGACGTGGACCGGCGGCTCGCGCGCGCGCGGCGTACGAGAAGGTTCGGGCGCGGATGCCGCGCGGCGGACCGGCGCAAGTACCTGTTC
 > A D V D R A V R A A R T A Y E K V W G P M P G R D R A K Y L F
 86173 CGGATCGCCCGATCATCCAGGAGCGCTCCCGCGAGCTGGCGGTGCTGAGTCCCTGGAACAACGCAACCGGAGTCCGGGAGTCCCGGAGCT
 > R I A R I I Q E R S R E L A V L E S L D N G K P I R E S R D V
 86265 CGACCTCGCGTGGTTCGCGCGGCGCTTCTTACTACGCGGCGTGGGCGAGCAAGCTGCGGTACGCGGCGTTCGCGCGGCAACCCCGCGGCG
 > D L P L V A A H F F Y A G G A D K L P Y A G F G P N P R P
 86357 TCGCGTGGCGCGCAGGTTCATCCCGTGAACCTTCCCGTGTCTCATGCTCGCTGGAAGATCGCCCGCGGCTGGCGCGCGGCAACACGGTG
 > L G V A A Q V I P W N F P L L M L A W K I A P A L A A G N T V
 86449 GTGCTCAAGCGCGGAGACACCGCGCTGCTGTTGCGCGAGATCTGCCAGCAGGCGGAGTGGCGGCGCGGCGTGGTCAACAT
 > V L K P A E T T P L T A L L F A E I C Q Q A E L P A G V V N I
 86541 CGTACCGCGCGCGGCGACACCGCGCGGCGGCTGGTTCGAGCACCAGCGGCGTGGACAAGGTTCGCGGCTTCACCGGCTCGACCGAGGTTCGGAAGG
 > V T G A G D T G R A L V E H P G V D K V A F T G S T E V G K
 86633 CCATCGCCCGGTTCGTCGCGGCGACCGGCAAGAAGGTACCCCTGGAGTGGGCGGCAAGGCGCGAATCGTCTTCGACGACGCGCGCGGT
 > A I R S V A G T G K K V T L E L G G K A A N I V F D D A P V
 86725 GACGAGGCGTTCGAGGGGATCGTCAACCGGCTGCTTCTCAACAGGGGCGAGTCTGCTGCGCGGCGTGGTTCGCGGAGGTTCGCTCGGAGTTCGCT
 > D Q A V E G I V N G I F F N Q G H V C C A G S R L L V Q E S V
 86817 CGCGGAGGAGTTCGAGTTCGTAAGCGCGGAGTGGCGGTGCTGCGGCTCGGCGACCCGTTGGACAAGAACACCGACATCGGCGGCGATCA
 > A E Q L K R E L R M A L V L G V G D P L D K N T D I G A I
 86909 ACTCGGCGCGGAGCTCGCGCGATCCGCGAGCTGTCCGCGCGGCGGAGGCGGAGGGGCGGAGCGCTGGTTCGCGCGCGTTCGCGGAGTTCGCG
 > N S A A Q Q L A R I R E L S A A G G E A E A E A E R W S P L P
 87001 GAGCGCGGTTCTGTTTCGCGCGGAGTTCACGCGGCGTACCGGCGGCGGAGGCGGAGGAGTTCGCGTTCGCGGCGTTCGCTGCT
 > E R G F W F A P T I F T G V T Q A H R I A R E E I F G P V L S
 87093 CGTGTGACCTTCCGACCCCGCGGAGGCGGCTCGAGAAGGCCAACACGCGGTACGGGCTGTGCGCGGAGTTCGACCGACAAGGGCT

Figure 11V

<G R H Y A S L L G R V G E P T R R Q V S T T P K E V G G P V L
 102426 TTGGCCATCGGCACCCGCTTCGGCAGGCCCTTCGGGTGGTCTCCGCCACCGGGACCTCGGTGGCCGCGAGGCGGCCCCGAGCCGCTCGT
 <Q A M P V R K P L G K R T T E A V P V E T A A S A A R W G D C
 102518 CCGCGGCAGTCTGCCAGGCGTGGGCTTCGGCGTGGGCGCGCGCGGCGAAGCCCTCGCGGGACCGGGCGGTTGCCACCGTTGGTGGC
 <A A T Q W A H A Q P T P R R G A F G E A P G P R T G G N T P
 102610 GACCCGTTGTGCGCGGCATTCTCCGCCATCGGGGTGTCTGCCATCGGTGCGTTACCTGTCTGCCCGGGTGTGGTGTCTGGACGGGCGG
 <S G N D R P M G G A M P T D A M P A N G T T G P A V P R
 102702 GCCGGTGACGTGACCGCGGAGAACTGTTGGGTACCGCGCGGCTTCGGCGGCGCATTGGTTCGCGCGCTGGGCGACCGCGCGG
 <G T V D V A S F Q Q T V A A N A P S G A G N T A R Q A V G A T
 102794 TCTCTCCGAACCCGAGCGCGGGTACGGAACAGCGCGACTCGAGCTCCCGGAAGATCGGCAGCTCCATCGTTCGTCGCGTACCGCTGC
 <E E S G S R R T R F W A S E L E R F I P L E M T E D A Y R Q
 102886 TGCCGTTCTGGGCTGACGCGCGCTGACCGCGCGCGCTGGCGGGGTCTCGGCTGGGTGGCGCGCTGCTCGGCACCTCGGTGCTCGG
 <Q R N Q A Q V P T S R A P T P T P T P Q T A P T S P V E T S P
 102978 CACCGGGGACCTCGGTGCTCATGTCCAGGCTGCGCGGAGCGCTCCGCCACCGCGGGGTGACCGGCTCCGCGCGCGCCACCGCGCGG
 <V R P L E T T M D L A A A L R E P V P P T V P E P A A V P P W
 103070 AGGCGCGCGCGCCACGGGCGCTGGGCGGACGGGACCGCGGATGGGACGCGGTGCGCGGAGTACGGCTGACCGGACCGCGCGGCGG
 <A P P A P A P A Q V P P R S P L P Q A S Y P Q G S V P T G
 103162 AACGGTGACCGGAGACGGGCGTGGCGAAGCGTACCGGACCGCGGGCGCGGAGACGGGTGACCGGAGACCGGGAAGACGGAACCGG
 <F P Q G S V P T G F P Q G S V P A A S V P Q G S V P P
 103254 CGGTGCGGACACCGCGGACCGGTGGCGGGTCCAGCCCCGGGCTCCGGGCTGCTCGGCAGCTGTGGGGATGGCGCGCTGCT
 <P A S S V P P V P P P T W G R A E P S S P L Q R P I A P Q Q
 103346 GGCGCTGCTGGCGGATCGCGCTGCGCGGACCGCGCGCTGCGCGGAGCGGTGCGTTCGACTGGCCATTGAGTTCGCGCGCGCGGCA
 <G S S A P D G D G S A R R Q P L P D S S Q G N S T R G A A A
 103438 CCGCCAGCGGTGCCACTGGCCCCGGTCAAGTCCGACAGCGCGCATCGACCGATGGAACCGGTGACGCGCGGCGTGGCGCGCTGGC
 <G A T G L D S W A P M S R M S G T S A P T S G H G N R
 103530 CGAGGCGGGTGAACGACCGCGCGCGGAGGTTACCTGGTTGCCGAGTGGCGCGCGCTGGGTGGGCGCGGTGTGCGCGGTTGTTGC
 <S A P D F S R G G L T V Q N G S H G P R Q T P A P T A P N G
 103622 CGAAGCCGCGAAGGACCGGCGCGGGGCGCGCGCGGCTGCGAGCTGGTGAGCGAGGCGGAGGCGCGGACCGCGCGCGGCTGC
 <F A A F A G L A P P A G G P Q S S T L S A P P A P L A G P Q
 103714 TGGAACCGCGGAGAGCGCCCCGGGACCGAGCAGGTGGTGGCGAGGTGACGTGCGCGACGGTCCCCCGGTGGTGGCGCGGCGAGCTC
 <Q F R G S L A R G V L V T T P L T V D V A V T G R D T G P R L E
 103806 GACCTTGACCCGCTGCGCGGACGCCAACCGGGCGACCAACAGGCGCATCATCGGGAGACGCCACGTCCACTGCGGCGCGGAGGCGA
 <V K V G H R S A L R A V V V L G M M R S V A V D V P Q A L
 103898 GGCGCTCGTTGAGTCTGTAGCTGCTGGCGCTGATGCGGATCGCCCGTCTCGACGTAGAGGTTGGCGCGGTGCGCGACCGCGCGGCG
 <R D N L D H L Q E A S I G I G R D E V Y L N A R D G V R R A
 103990 TCCACCATCACCTGCGAGTGGGCGGCGAGAAGCGGTGCGGTGTCGAACAGCTCGCGGACCGAGTGGACCGAGTGGTACCGCGTGC
 <E V M V Q S D P P S F A T A N D F L E A V L H V L D N V A H A
 104082 GGCGACCTCGATGTCACGGTTCGATCACCGGAACTCGATCGGGTGTAGTGTGACCTCGGACTGGGCGCGCGCGAGCAGCTCGATCAGT
 <A V E I D R D I V G F E I R T Y H E V E S Q A A R L D I L A
 104174 CCGCGCGCTGCGGCTGACCGCGGCGGAGTGGAGTGGCGCGGAGCAGGAGGTTCTCGTCTGCGCGCATCCGGGTGGCCAGGTGGTGC
 <A P E R Q V R T S D A G A L V L L N E D N R R M R T A L H D
 104266 AGCTGGAACAGCTCGGCCAGCGGTCCCTCTCGCGCGCTCCAGCGGTGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGA
 <L Q F G A L R D P D E G R E L R D L H G I L R D V L I Q S
 104358 ACGGCGGGCAGGTTGACGAACATGGTTCGCGCAGGAGGCGCGCAGCGCGGCTGCTCGCGCGCGGTCGTAAGCGCTCCAGGTGGACCGGT
 <R R A L N V F M T A V S R A L A A Q E A A T R V A E L H V A N
 104450 TGAACGCTCGGTACCTGGCCGAACCTCGTCTTGTGCGCACCGGCGAGCGGCTCGGCGATCTGGTTGGCGCGCTGACCGGGGAGAGCTGG
 <F A E T V Q G F E D K S R V P L P E A I Q N A A Q V P S L Q
 104542 CTGGAAAACTCGGGATCGCGCGAGCGGCAAGCGCTGGGCGAAGCGTGGGCGATGCTGAGCGCACCTGGCGGAGGTGCGCGAGCGA
 <S S F Q P D R L R A V A Q P L G Y Q A I S L A G Q R L D R L S
 104634 GCGGCGCATCGACCGGCGACCGAGTACCGGAACAGGATGGCCAGGAGCAGCATGCGGAGCAGCAGGCGCGGTCTGGAGGAACACCGGTGCGCT
 <R A M L A V A F L I A L L M G L L L G L L G Q L Y V T R Q
 104726 GTACGTGCGGAGCGGAGCGCTGCGGCTGCTTGACACGTTGCGCTCGAGCTTCGCTCGACCGTACGGATCAGTTTGGCGCTGGCGACCATG
 <V D S R L A D A Q K V V N G D L K A E V T R I L K A S A V M
 104818 GCCGCTCCCACTGACCGCGCGGCGGCGGCTGCGGCTGCTGCGGTTGCGGCTGCGGAGCGCGGTGAGTTCTGCGCTGCGG
 <A A D W Q D P G F P A N A M S G N T N G D L W G T Y N Q A E R
 104910 CCGGTGCGCGCGCGCGCGGCTGCTGGTGTGCGAGGTCCGACTCGTCCAGGCTGGCCACCGCTTGAAGCTTGTGAGTGTGCTGCTGCGCGG
 <R D G G A V T D D H L D S E D L S A V A K F S Q L A Q Q G T
 105002 TGCCGCTGCGCATGTAGTGGTTCGCGCAGGATGGGGTCAACTCGCGTGGATCAGCGCGCGGTGCGACCGCGCGGCGGAGAGGTAT
 <G S A I Y D T R L I P T L E R Q I L A R H V V V R S L Y
 105094 TCCTTCTCCCGGCGGCGGCTGCGCGGCGCGCATCCGCTCGCTCAGTCTGCTGTCACCGGCGAGGTGGGTGGCGGAGTTCGCGGATGGACAG
 <E K E R A V A A A A R M R D S L D N D G A L H T A S D R I S L
 105186 CAGGTCTGATCAGGCGCTCGTACGCTGCGGCTGATGATCTCAACTTGGCGTTGAAGACCTGGCTGGCGGTGCGCGGAGGTCTCT
 <L D N I L G E Y A Q M A D I I K L K G N F V Q S R T G P L D K
 105278 TCAGGTCTGCTGATCGCGTGGAGGCGGCTCCAGGCTGCTCGCGAGGCGGCTCGACCTCCCCCGCTGTTGAGGTACGGCACCTTGTCTC
 <L N Q D I G D L L G L L S P L G D V E G R Q Q L Y V K D
 105370 TGGTCGACCGGATGTTGACCGGTTGTACGCTCTCTGTTACTGCGCTTGGGCTGGTTCGCGCTGCGCGCGGAGCAGCAACACCGCGGAGGT
 <Q D V R I N V R N Y A E Q Y Q A K A Q D G S A G L L L V A S T
 105462 GCGTTCGCTTGGAGGCTGTTGACCGGAGTACCGGAGTGGCGAGGTGCGCGGAGCGGTTGGCGTTGTTGAGCGTTTTCAGCGTTTCCA
 <R E D Q L S N V L D G S Y G V L N A L D G S R N A N N L T E L
 105554 GGTGTCGACGAGGCGGCTGGTGGCGGACCGGCGGATGCTGCGGACGATCATGATGAGACCGAGCTTGGACGAGTGGACGAGTGGCATGTC
 <N D V L G S T G G V V V T A I T P V I M I L G L K S W I P M
 105645 GCGGAGCGCGCGACCGGCGGACGAGTGGCGACGGAAGAACCGCGGCTCTCGGTGCTTGTCTACGTCACCGCGCTCGCGATCACAGC
 105737 GTTCGCGGCTGCGCGGCGCAAGCTCAGCGACCGACCGCGGCGGTTCGAGCTTCGAGATTCCATCAGCGCGGTGTTCCAAAGAGAAAGCCCA
 105829 GGCTGGCGCTGCGCGGAGGTGTGATGAGATGTTGATGCAATTTGCTCGCAATCCGTCACGCGGAGTGACTGACGAGTACCGCTCAACCC
 >M D H P
 105919 CACCGCTCGTCTGCTCGCGGCGCTTCGGGCTCCGGAAGGCTGATACATAGCCCAACAAACCGGCTTCTGTTCTGTCGACGACTT
 >H R L A L G P S G S G K S Y I A Q Q T G L P V L L G L D D F
 106011 CTACAAGGATGGTATGACCTACGTTACCGCGCAAAACGGTCTGTGGACTGGGACTACCCCACTCGTGGGACCGCGGCGGCGCGGTGG
 >Y K D G D D P T L P R G N G L V D W D S P Q S W A P V
 106103 AAACGATGCGCGGCTGGCGGCGGACGCGCAAGGCGAAGTTCGCGTTATGCGATCGGCGGAGCGGCGGTTGGCCACCGGACATTTCGAG
 >E T I A R L A R D G K A E V P V Y A I G A D R R V A T R T F E
 106195 GTCGCGGATCGCCACTTTTCGTCGCGCAAGGATTTTCGCGCGGAGATCGTCGAGGAATGCCGACGCGGAGGCTGCTGCGCGGCGGCTA

Figure 11AA

1 GTCTTCGGGGAACGCACGGGAACCTCTCTCGGCAGAACCGGACGAGCTGACCGTAACCGGTGGGACACCCGGCGGGATCAACCGGTGCGGTTCACAGGTTCCGGCCGGCGGCCCGG
122 ACGCACCGCCCCCGGTGTGAGAGGGGACCCATCTCTACCGTAGGCGTTAAACAAGGGCCCTTCCTT TCA CCG CAG GTG CAG GAC GTC TCC TGC GGC GAG GGT GCG
228 CGG GCC CGC CGG GGT GTC GAC CAG CAG CCG GCC GTC GGC GTC CAC GCC GGT CAC CTC CGC ACC GCC GGC CAG CAG CAG CAC
< P G A R T D V L R G D A D V G T A V G T A V G T V E A G G P L L V
318 CCG CAC CGG TCT GCC GAC CGT CGC GCA GGC CGC CAG GTA CGC GTC GCG CAG CCC GCT GGC CAC CGC GTC GCC GGC GTG GCG CCA CCG
< R V P R G V T A C A L Y A D R L G S A V A D G G A H R W R
408 GTC GTA CCA GTC GGC GAC AGA GCG CAG CAG TGC GCG CAG CAG CGG GTC CCG GTC GGT GGC GAC CCC GGC GAG CTG CAG CAG GGT GGC
< D Y W D A V S R L L A R L L P D R D T A V A G A L Q L S T A
498 CCG CAG GCC GGT CGG GTT CGC CGG CAG CTC GTC GGC CGC CAG GGT GAC GTT GAT ACC GAT GCC GAG GAC GAT CGC CGG GGC GTG GTC CCG
< P L G T P N A P L E D A R L T V N L G I G L V I A P P Q D P
588 GGC CCG GCC GGC CAC CGC CTC GGC CAG GAT GCC GCA CTT GGC GTC GCC GAT CAG CAG GTC GTT GGC CCA CTT GAG GGC GGC GTC CAG
< A P G P V A E A L I G A C K A D G I L L D N P W K L A A D L
678 CTC GGC CAG CGC GGC CAC CGC CTC GAC CAG CGC GAC GCC GGC GAG CAG GGC CAG CCA GCC GTA CCC GGT CGC GGC GGC CGG CCA GTC
< E A L R A V A E V L A V G A L L P L W G Y G T A P A P P W D
768 GCG CTC GCG GAC AGC CTC GCC CGC CAG CAG GAT GGT GGT CGC GCG GCG CCA GAT CGC CCA GAC CCG GCC CCG CCG GCC
< R E A V A E G P R L L V S T A I G A R P P S Q W V R G R G
858 CCG GCC GGT CTG CCG CTC GGC GAT CAC CAG GAC GCC CTC CGC CTC GGC GGA TCG GGC CGC CTC CGC CAC GTC CGC GTT GGT CGA GCC
< R G A T Q R E A I V V L G E P E G S R A A E A V D A N T S G
948 GGT CTC GGC GCG TAG CTC CAG CCG GCC CCA GGC GCG GGT CAG CGC CCG GCG CAG CCG GGC CGC CGA CAG CAG CCG GCG ATC
< T E A R L E L R A W P G H P A T L A R R L R A A S L P P R D
1038 CAG GTC GGT GTA CCG CGA GCC GGC CAT CCCGCCAGCCTACGGCCCGCCCGGCGGTGCGCGGTTGCCCGAGCCCGGTGAGTTGCTGA GGC GTA CTG
< L D T Y P S G P M
1145 CAC ACC GTC GGC CAC CTG AAC CAT CGT TAT ATT CCG TGG GTG ACT ACC GAG ACC GGG ATC AAC ATC CAC AGC ACG GCG GGC AAG CTG GCG
1235 GAC CTG GAG CGA CCG GTC GAC GAG GCG GTG CAC GCC GGA TCG GCG CGT TCC AAG CAG CAC GCC CGG GGC AAG AAG ACG GCG CGG
> D L E R V D E A V H A G S A R A V S K Q H A R G K K T A R
1325 GAG CCG ATC GGC CTG CTG CAC GAG GGC TCC TTC GTC GAG CTG GAC GGG TTC GCC CGG CAC CGG TCC ACC AAC TTC GGC CTG GAC CGC
> E R I G L L L D E G S F V E L D G F A R H R S T N F G L D R
1415 ACC CCG TAC GGC GAC GGC GTG ATC ACC GGC TAC GGC ACG GTC GAC GGC CGG CAG GTC TGC GTC TTC GCG CAG GAC TTC ACG GTC TTC
> T R P Y G D G V I T G Y G T V D G R Q V C V F A Q D F T V F
1505 GGC GGC TCC CTC GGC GAG GTG TTC GGC GAA AAG ATC GTC AAG GTG ATG GAC CTG GCC ATG AAG ATC GGC TGC CCG GTC GTC GGC ATC AAC
> G G S L G E V F G E K I V K V M D L A M K I G C P V V G I N
1595 GAC TCC GGC GGC CTC CAG GAG GGC GTG GCG TCC CTC GGC GAT ATC TTC TGC AAC GTG CCG GGC GGC AGC GGC GTC
> D S G G A R I Q E G V A S L G L Y G E I F R N V R A S G V
1685 ATC CCG CAG ATC TCC CTG ATC ATG GGC CCG TGC GCG GGC GAT TAT TCT CCG GCG GTC ACC GAC TTC ACC GTG ATG GTC GAC CAG
> I P Q I S L I M G P C A G G A V Y S P A V T D F T V M V D Q

FIGURE 12A

1775	ACC	TCG	CAC	ATG	TTC	ATC	ACC	GGC	CCC	GAC	GTG	ATC	AAG	ACG	GTC	ACC	GGC	GAG	GAC	GTC	GGG	ATG	GAG	GAA	CTG	GGC	GGT	GCC	CGC	ACC
>T	S	H	M	F	I	T	G	P	D	T	V	I	K	T	V	T	G	E	D	V	G	M	E	E	L	G	G	A	R	T
1865	CAC	AAC	CGC	AGC	GGC	AGC	GGC	ACC	GGC	TAC	CTC	GGC	ACC	GAC	GAG	GAG	GAC	GCG	ATC	GAG	TAC	GTC	AAG	CGC	CTG	CTG	TCG	TAC	CTG	CCG
>H	N	A	R	S	G	N	A	H	Y	L	G	T	D	E	E	D	A	I	E	Y	V	K	A	L	L	S	Y	L	P	
1955	TCG	AAC	AAC	CTG	GAC	GAG	CCG	CCG	CTC	TTC	GAC	GCC	CCG	GCC	GAC	GTG	GCG	ATC	AGC	GAC	GCC	GAC	CGG	GAG	CTG	GAC	AGC	CTC	GTC	CCG
>S	N	N	L	D	E	P	P	V	F	D	A	P	A	D	V	A	I	S	D	A	D	R	E	L	D	S	L	V	P	
2045	GAC	TCG	GCG	AAC	CAG	CCG	TAC	GAC	ATG	CAC	CCG	GTC	ATC	GAG	CAC	GTG	CTG	GAC	GAG	GGG	GAG	TTC	CTG	GAG	GTC	CAG	CCG	CTG	TAC	GCG
>D	S	A	N	Q	P	Y	D	M	H	R	V	I	E	H	V	L	D	D	G	E	F	L	E	V	Q	P	L	Y	A	
2135	CAG	AAC	ATG	GTG	GTC	GGC	TTC	GGT	CGA	ATC	GAG	GGG	CGA	CCG	GTC	GGC	GTG	GTG	GCC	AAC	CAG	CCG	ATG	CAC	CTC	GCC	GGC	ACG	CTG	GAC
>Q	N	M	V	V	G	F	G	R	I	E	G	R	P	V	G	V	V	A	N	Q	P	M	H	L	A	G	T	L	D	
2225	ATC	GCC	GCG	TCG	GAG	AAG	GCC	GCC	CTG	TTC	GTG	CGC	ACC	TGC	GAC	GCG	TTC	AAC	ATC	CCC	GTG	CTG	ACC	TTC	GTG	GAC	GTG	CCC	GGG	TTC
>I	A	A	S	E	K	A	A	R	F	V	R	T	C	D	A	F	N	I	P	V	L	T	F	V	D	V	P	G	F	
2315	CTA	CCC	GGC	ACC	GGC	CAG	GAG	TGG	GAC	GGC	ATC	ATC	CGG	CGC	GGC	GCG	AAG	CTC	ATC	TAC	CGC	TAC	GCC	GAG	GCG	ACC	GTC	CCG	AAG	GTC
>L	P	G	T	G	Q	E	W	D	G	I	I	R	G	A	K	L	I	Y	A	Y	A	Y	A	E	A	T	V	P	K	V
2405	ACC	GTG	ATC	ACC	CGC	AAG	CGC	TAC	GGC	GGG	CGC	TAC	GAC	GTG	ATG	GGC	TCC	AAG	CAC	CTG	GGC	GCG	GAT	CTG	AAC	TTC	GCC	TGG	CCG	ACC
>T	V	I	T	R	K	A	Y	G	G	A	Y	D	V	M	G	S	K	H	L	G	A	D	L	N	F	A	W	P	T	
2495	GCG	CAG	ATC	CGC	GTG	ATG	GGC	GCG	CAG	GGC	GCG	GTG	AAC	ATC	CTG	TAC	CGG	CAG	GAG	CTG	GCC	GCC	GAG	GAC	C'CG	GCC	GCC	GTG	GTC	GCG
>A	Q	I	A	V	M	G	A	Q	G	A	V	N	I	L	Y	R	Q	E	L	A	A	A	E	D	P	A	A	V	R	
2585	GCC	GAG	AAG	ATC	GCC	TAC	GAG	GAC	ACC	CTG	GCC	AAC	CCG	TAC	GT'C	GCC	GCC	GAG	CGC	GGG	TAC	GTC	GAC	TCG	GTG	ATC	CCG	CCG	CAC	
>A	E	K	I	A	E	Y	E	D	T	L	A	N	P	Y	V	A	A	E	R	G	Y	V	D	S	V	I	P	P	H	
2675	GAG	ACG	CGT	ACC	CAG	ATC	GT'C	CGG	CGC	TTG	CGG	GTG	CTG	CGC	ACC	AAG	CGC	GAG	ACG	CTC	CCG	CGG	AAG	CAC	GGC	AAC	ATC	CCG	CTC	
>E	T	R	T	Q	I	V	R	A	L	R	V	L	R	T	K	R	E	T	L	P	A	K	K	H	G	N	I	P	L	
2765	TAG	GGCGGTG</																												

FIGURE 12B

FIGURE 12C

FIGURE 12E

FIGURE 12F

[illegible]

FIGURE 12G